



UNIVERSIDAD EAFIT

TESIS DE MAESTRÍA

Modelo matemático combinado para la clasificación de neuroimágenes basado en medidas de similaridad entre hemisferios del cerebro

Autor:

Danny Styvens CARDONA
PINEDA

Asesor:

Henry LANIADO RODAS

Co-Asesor:

Juan Gabriel CASTRILLÓN
GUZMÁN

*Trabajo para optar al título de Magíster en
Ciencia de Datos y Analítica*

En las facultades de

Ciencias, Ingeniería y Economía

julio 22, 2020

Resumen

Partiendo de la hipótesis de simetría entre los hemisferios del cerebro de una persona, se estudiaron tres medidas de similaridad para diferenciar entre un cerebro sano y uno patológico a partir de neuroimágenes. Se usaron tres medidas, la primera es un índice de correlación cruzada (NCC) [51], el segundo es una prueba basada en permutaciones de valores de intensidades entre hemisferios (HCP) y el tercero es la medición de la entropía de valores locales permutados entre hemisferios (HPLE). También se calcularon mapas de probabilidad con el método BIANCA [19] para usarlo como entrada de las medidas de similaridad.

Se observó que el HCP y el HPLE (con un 5 % de significancia) obtuvieron valores del *accuracy* iguales al 66 %, sin embargo, la capacidad de detectar cerebros patológicos fue mayor para el HPLE. Los resultados obtenidos con estas dos medidas no fueron lo suficientemente satisfactorios para sugerir que se pueden usar en el ámbito clínico, aunque el HPLE puede ser mejorado.

El uso del mapa de correlaciones NCC calculado a partir de una secuencia FLAIR para medir similaridad entre hemisferios, permitió la clasificación de cerebros sanos y patológicos mediante la generación de un subespacio con el método Eigenfaces y aplicando modelos de clasificación a los datos proyectados al espacio propio hallado. El proceso de reproducibilidad incluido en la generación de espacios propios y el ajuste de los modelos de clasificación, permitió atribuir los resultados al método y no a la elección de un subespacio específico. Se observó para esta combinación de métodos unos resultados con un *accuracy* del 81 % y una capacidad predictiva del 79 %.

La principal contribución de este trabajo es la combinación de medidas de similaridad, métodos para la construcción de subespacios y modelos de clasificación. Específicamente, se usó el NCC como medida de similaridad, la cual fue proyectada al espacio propio en descomposición de valores singulares siguiendo la metodología Eigenfaces [58], para luego aplicar modelos de clasificación sobre estas proyecciones.

Palabras claves: Mapa de correlaciones, BIANCA, permutaciones, entropía, Eigenfaces, modelos de clasificación.

Abstract

The healthy brain is highly symmetric, therefore, we proposed to use similarity measures in neuroimaging data to distinguish between a healthy and a pathological brain. Three measures were used, the first one Normal Cross Correlation (NCC) [51], the second one is a test based on permutations of values of intensities between hemispheres (HCP) and the third one is a measurement of entropy of values local exchanged between hemispheres (HPLE). Furthermore, we also used the maps derived from BIANCA as an input for the similarity measures [19]

HPLE(with 5 % significance) provided better results compared to the HCP measure both obtained *accuracy* values equal to 66 %, however, the ability to detect pathological brains was greater for HPLE. However, HPLE is not good enough for clinical applications.

The use of NCC correlation map derived from FLAIR images to measure similarity between hemispheres allowed the classification between healthy and pathological brains by generating a subspace with the Eigenfaces method and applying classification models to the data projected to the proper space found. The reproducibility process included in the generation of own spaces and the adjustment of the classification models, allowed to attribute the results to the method and not to the choice of a specific subspace. Results were observed for this combination of methods with an *accuracy* of 81 % and predictive capacity of 79 %.

The main contribution of this project is the combination of similarity measures, methods for the construction of subspaces, and classification models. Specifically, the NCC was used as a measure of similarity, which was projected into the own space of decomposition of singular values following the Eigenfaces methodology [58], finally, classification models were fitted.

Keywords: Correlation map, Bianca, permutations, entropy, Eigenfaces, classification model.

Modelo matemático combinado para la clasificación de neuroimágenes basado en medidas de similaridad entre hemisferios del cerebro

por

Danny Styvens CARDONA PINEDA

Estadístico *Univerdiad Nacional de Colombia* MSc Epidemiología *Univerdiad de
Antioquia*

Trabajo para optar al título de magíster en
Ciencia de Datos y Analítica

en

Universidad EAFIT

julio, 2020

Declaración

Yo, **Danny Styvens Cardona Pineda**, declaro que esta tesis titulada, «Modelo matemático combinado para la clasificación de neuroimágenes basado en medidas de similaridad entre hemisferios del cerebro», que se presenta en cumplimiento de los requisitos para el grado de Magíster en Ciencia de Datos y Analítica, representa mi propio trabajo, excepto donde se haya hecho el debido reconocimiento. Además, declaro que no se había incluido previamente en una tesis o informe presentado a esta Universidad o cualquier otra institución para obtener un título, diploma u otras calificaciones.

Danny Cardona P.

Firma: _____

Fecha: julio 22, 2020

Dedicatoria

El presente trabajo está dedicado a mi familia por su apoyo incondicional. A todas las personas especiales que me acompañaron en esta etapa, aportando a mi formación profesional y como ser humano.

Agradecimientos

En primer lugar deseo expresar mi agradecimiento a los asesores de este trabajo de grado, Dr. Henry Laniado Rodas y Dr. Gabriel Castrillón Guzmán, por la amable dedicación y apoyo durante este proceso. Agradezco a colegas y compañeros de trabajo quienes me hicieron valiosos aportes, especialmente Catalina Bustamante, Luisa Sanchez y Aura Puche, con quienes he tenido la fortuna de trabajar en diferentes proyectos y me orientaron en temas relacionados con el procesamiento y análisis de neuroimágenes. También agradezco a los gerentes Regina Sepulveda Montoya y David Sanguino Cotte, quienes además de brindarme su amistad, me apoyaron en representación de Servicios de Salud - San Vicente Fundación.

Gracias a mi familia, mi esposa y padres por toda su paciencia, comprensión, apoyo y cariño. Gracias a mis amigos, que siempre me han prestado un gran apoyo moral. A todos, muchas gracias.

Danny Styvens CARDONA PINEDA
julio 22, 2020

Índice general

Resumen	I
Abstract	III
Declaración	I
Agradecimientos	II
Índice de figuras	V
Índice de cuadros	VII
Lista de abreviaturas	IX
1. Introducción	1
2. Marco teórico	5
2.1. Resonancia magnética y neuroimágenes	5
2.2. Simetría entre hemisferios del cerebro humano	6
2.3. Coeficiente de correlación cruzada (NCC)	7
2.4. BIANCA - <i>Brain Intensity Abnormality Classification Algorithm</i>	8
2.5. Permutaciones de vóxeles entre hemisferios (HCP)	9
2.6. Entropía de valores locales permutados entre hemisferios (HPLE)	10
2.7. Eigenfaces	12
2.7.1. Análisis de componentes principales (PCA)	12
2.7.2. Eigenfaces: PCA modificado	13
2.8. Modelos de clasificación supervisados	15
3. Objetivos	19
3.1. General	19
3.2. Específicos	19
4. Métodos	21
4.1. Tipo de estudio	21

4.2. Población	21
4.3. Muestreo	21
4.4. Tamaño de la muestra	21
4.5. Selección de la muestra	22
4.6. Criterios de inclusión	22
4.7. Criterios de exclusión	22
4.8. Fuentes de información	22
4.9. Control de sesgos y overfitting	22
4.10. Limitaciones del estudio	23
4.11. Procesamiento de imágenes	23
4.12. Preparación y partición de datos para el ajuste de modelos	24
4.13. Variables en el estudio	24
4.14. Consideraciones éticas	24
5. Resultados	27
5.1. HCP	28
5.1.1. HCP con valores de intensidades normalizados	29
5.1.2. HCP con mapas de probabilidad	31
5.2. HPLE	33
5.3. Eigenfaces	35
5.3.1. Eigenfaces por hemisferios y grupo con valores de intensidades en $[0, 1]$: Enfoque A	37
5.3.2. Eigenfaces con NCC : Enfoque B	38
5.3.3. Eigenfaces con BIANCA : Enfoque B	41
6. Discusión y conclusiones	43
6.1. Discusión	43
6.2. Conclusiones	45
A. Propuesta de otras líneas de trabajo relacionadas	47
B. Código	49
Bibliografía	51

Índice de figuras

2.1. Coeficiente de correlación cruzada - NCC	7
2.2. Mapa de probabilidades por hemisferio obtenidos con BIANCA, en un sujeto diagnosticado con oligodendroglioma	9
2.3. Selección de regiones en ambos hemisferios para calculo del HPLE, sin hacer permutaciones	11
2.4. Selección de regiones en ambos hemisferios para calculo del HPLE, haciendo permutaciones	11
4.1. Procesamiento de neuroimágenes de RM.	23
5.1. Distribución de las diferencias entre hemisferios ajustadas, según el grupo.	28
5.2. Distribución del test HCP en 20 neuroimágenes de sujetos sanos.	29
5.3. Distribución del test HCP en 20 neuroimágenes de sujetos pato- lógicos.	30
5.4. Distribución del test HCP en 20 neuroimágenes de sujetos sanos - BIANCA.	32
5.5. Distribución del test HCP en 20 neuroimágenes de sujetos pato- lógicos - BIANCA.	33
5.6. Distribución del HPLE entre hemisferios, según el grupo.	34
5.7. Tendencia de indicadores de desempeño del HPLE según el nivel de significancia, aplicando prueba U de Mann-Whitney.	34
5.8. Tendencia de indicadores de desempeño del HPLE según el nivel de significancia, aplicando prueba t de student.	35
5.9. Diferencia entre distancia al respectivo espacio propio de cada hemisferios respectivamente.	37
5.10. Validación cruzada de modelos de clasificación, Eigenfaces enfo- que A.	38
5.11. Validación cruzada de modelos de clasificación, Eigenfaces enfo- que B.	39
5.12. Capacidad predictiva del modelo SGB usando un espacio propio que retiene el 90 % de la variabilidad, Eigenfaces enfoque B.	40

5.13. Validación cruzada de modelos de clasificación, Eigenfaces enfoque B con BIANCA.	41
5.14. Capacidad predictiva del modelo SGB usando un espacio propio que retiene el 90 % de la variabilidad, Eigenfaces enfoque B con Bianca.	42

Índice de cuadros

4.1. Variables en el estudio	24
5.1. Medidas de resumen del test HCP de sanos y patológicos.	29
5.2. Clasificación según la distribución del test HCP	31
5.3. Medidas de resumen del test HCP de sanos y patológicos - BIAN- CA.	32
5.4. Clasificación según la distribución del test HCP	32
5.5. Capacidad predictiva del HPLE según niveles de significancia, aplicando prueba U de Mann-Whitney.	35
5.6. Capacidad predictiva del HPLE según niveles de significancia, aplicando la prueba t de student.	36
5.7. Desempeño del modelo SVM, Eigenfaces enfoque A.	38
5.8. Desempeño de modelos sobre datos de prueba, Eigenfaces enfo- que B.	40
5.9. Desempeño del modelo SGB usando un espacio propio que retie- ne el 90 % de la variabilidad, Eigenfaces enfoque B.	40
5.10. Desempeño de modelos sobre datos de prueba, Eigenfaces enfo- que B con BIANCA.	41
5.11. Desempeño del modelo SGB usando un espacio propio que retie- ne el 90 % de la variabilidad, Eigenfaces enfoque B.	42

Lista de abreviaturas

BIANCA	Brain Intensity Classification Algorithm
PCA	Principal Components Analysis
HCP	Hemisphere Cross Permuted
HPLE	Hemisphere Permuted Local Entropy
SVM	Support Vector Machine
NB	Naive Bayes
RF	Random Forest
CART	Classification And Regression Trees
LR	Logistic Regression
AGB	Adaboost Classifier
SGD	Stochastic Gradient Boosting
FLAIR	Fluid Attenuation Inversion Recovery
PAC	Probably Approximately Correct Learning

Capítulo 1

Introducción

El cerebro humano es un universo fascinante que esta siendo explorado hace décadas, en esta exploración, personas con una gran visión y duro trabajo han logrado descifrar aspectos importante sobre funciones relacionadas con regiones específicas y detalles anatómicos de gran impacto para la ciencia [51, 2]. Aun así, todavía falta mucho por descubrir y aprender sobre este excepcional órgano.

Gracias al desarrollo de tecnologías como la resonancia magnética (RM) se tienen nuevas opciones para la investigación y el diagnóstico en el área de neurociencias. Una de las ventajas de la RM comparado con otras tecnologías como la tomografía computarizada es que no se expone a los pacientes a radiación ionizante, además de su utilidad clínica como ayuda diagnóstica.

La posibilidad de cuantificar y clasificar características del cerebro humano mediante técnicas no invasivos como la RM [59], justifica y le da importancia a la búsqueda de métodos y métricas que permitan detectar hallazgos útiles para la practica clínica [31, 12]. Son de gran interés este tipo de esfuerzos que buscan ser un apoyo en el diagnóstico que realizan los radiólogos, especialmente en enfermedades de alto impacto en la sociedad como el Alzheimer[31, 42, 37] o tumores cerebrales [29]. El uso de métodos estadísticos y herramientas como el *machine learning* hacen parte de las opciones con las que se tratan de clasificar neuroimágenes a partir de la extracción de características del cerebro que sean relevante para el diagnóstico de enfermedades [42, 37, 35, 5].

Partiendo del conocimiento anatómico del cerebro humano, específicamente del concepto de simetría aproximada entre hemisferios, en este proyecto se hace una búsqueda de características cuantitativos en neuroimágenes de RM, las cuales permitan la clasificación de sujetos sanos y enfermos (tumores, eventos cerebrovasculares, enfermedad de sustancia blanca).

Bajo la hipótesis de simetría entre los hemisferios del cerebro, se estudiaron medidas de similaridad para la detección entre un cerebro sano y uno patológico. Se estudiaron tres medias, la primera es un índice de correlación cruzada (NCC) [20] con el que se construyó un mapa de correlaciones, el segundo es una prueba basada en permutaciones de intensidades entre hemisferios (HCP) y el tercero se trata de la medición de la entropía de valores locales permutados entre hemisferios (HPLE). También se calcularon mapas de probabilidades con un método llamado BIANCA [19] para usarlo como entrada de las medidas de similaridad, y se compararon con los resultados obtenidos con las intensidades extraídas de la secuencia de recuperación de la inversión atenuada de fluido (FLAIR).

Con el fin de probar si la extracción de características adicionales a las consideradas con el HCP y HPLE, permiten obtener mejores resultados, se aplicó la técnica Eigenfaces [58] sobre el mapa NCC y el mapa de probabilidades obtenido con BIANCA. Con este método se obtuvieron subespacios con las principales características del conjunto de imágenes de sujetos sanos. Finalmente, se ajustaron modelos de clasificación para la predicción de las categorías sano/patológico.

La principal contribución de este trabajo es la combinación de tres métodos para clasificar las neuroimágenes, esta combinación consistió en la proyección del mapa de correlación NCC sobre el espacio propio en descomposición de valores singulares siguiendo la metodología Eigenfaces [58] y ajustando modelos de clasificación sobre estas proyecciones para distinguir entre un cerebro sano y uno patológico.

Es importante mencionar que previo a la aplicación de métodos y pruebas que se acaban de mencionar se hizo un procesamiento de las imágenes DICOM. El procesamiento consistió en la extracción del cerebro, corrección por movimientos y normalización al espacio estándar MNI 152 [50, 8], además se extrajo cada hemisferio del cerebro.

Este documento se divide en las siguientes secciones: en el capítulo 2 se presenta una descripción de los temas más relevantes en el proyecto, tales como resonancia magnética en neuroimágenes, índice de correlación cruzada, BIANCA, pruebas basadas en permutaciones y medidas de incertidumbre, Eigenfaces y modelos de clasificación; en el capítulo 3 se encuentran los objetivos; en el 4 los métodos incluyendo las consideraciones éticas; en el 5 los resultados y en el 6 la discusión y conclusiones. También se incluyeron en los anexo A y B el

código usado y propuestas de otras líneas de trabajo, además de las referencias bibliográficas.

Capítulo 2

Marco teórico

En este capítulo se hace una breve descripción de los temas más relevantes en este proyecto, para una mayor profundización se recomienda al lector revisar las referencias suministradas en cada uno de los temas.

2.1. Resonancia magnética y neuroimágenes

Las imágenes de resonancia magnética tal como se conocen hoy tienen sus fundamentos en la resonancia magnética nuclear. En términos generales lo que pasa en el proceso es que a partir de la generación de un campo magnético de gran intensidad y señales externas al campo se excitan ciertos átomos en el organismo los cuales absorben y emiten energía, luego las señales de excitación pueden ser procesadas para obtener imágenes de la estructura del cuerpo estudiada. Hay dos conceptos importantes en resonancia magnética; el tiempo de relajación característico (T1) y el tiempo característico constante (T2) [34, 39].

Las neuroimágenes son especialmente relevantes dentro de la RM porque han permitido estudiar de manera más amplia el cerebro humano, estas se dividen en dos tipos: estructurales y funcionales. Con las estructurales se busca analizar el cerebro desde un punto de vista anatómico, mientras que en el segundo tipo el foco es la funcionalidad (áreas o redes de activación y flujos) [54]. Dentro de las patologías o hallazgos en las que se usan las neuroimágenes estructurales como ayuda diagnóstica se encuentran los tumores, epilepsia, hemorragia cerebral, enfermedades neurodegenerativas, entre otras [54, 45]. Por su parte la resonancia magnética funcional (fMRI) ha permitido detectar posibles áreas o redes afectadas por una patología o traumas que están asociadas a una función específica como el lenguaje o el movimiento [52, 43].

2.2. Simetría entre hemisferios del cerebro humano

Algunos estudios han abordado la detección de tumores y otras lesiones mediante segmentación y predicción automática [53]. En este trabajo se incluyen conocimientos anatómico en los análisis [38], con lo que se busca construir modelos matemáticos para la detección de imágenes sospechosas, una vez detectadas, se podría hacer una priorización automática dentro de los sistemas de información para que sea vista rápidamente por el personal médico. Se propone medir la simetría entre hemisferios como indicador de cambios anatómicos sospechosos de lesiones o trastornos neurológicos [38, 30, 56].

Algunos autores analizan la asimetría entre estructuras de hemisferios del cerebro en términos de dimensiones (largo, ancho, alto y radios) encontrando posibles diferencias, aunque estos resultados no logran ser consistentes aun [60]. Sin embargo, en este trabajo se propone abordar el problema como una relación entre los vóxeles de cada hemisferio, o en su defecto, entre regiones locales, pero en ambos casos manteniendo la relación espacial como factor de comparación, es decir que se hace necesario hacer un procesamiento de las imágenes que permita sobreponer un hemisferio sobre otro (*flip*) para asegurar que se está en un mismo espacio. Por supuesto, llevando antes las imágenes a lo que se conoce en neuroimágenes como un espacio estándar para que puedan ser comparados los resultados entre imágenes de cerebros que provienen de distintos sujetos [16].

Dado que las imágenes obtenidas con RM no solo incluyen tejido cerebral y que existen diferencias anatómicas entre cerebros de diferentes sujetos, es necesario realizar una extracción sobre la imagen original [50] y una normalización que permita hacer comparaciones entre cerebros de varios individuos [8]. Después de realizar esta normalización las relaciones de simetría entre hemisferios podrían variar entre cerebros. Por este motivo se propone una medida “intrínseca” de cada cerebro. Básicamente se busca conocer cual es la relación entre valores de vóxeles entre hemisferios de cada individuo para tratar de identificar anomalías. Note que bajo este enfoque podrían existir algunas variaciones entre hemisferios que serían biológicamente plausibles a partir de las mediciones que se le hagan a los sujetos sanos, aun así, las variaciones en esta medida no es un motivo de preocupación en esta investigación ya que lo que interesa es la relación intrínseca, en resumen, lo que se propone es tratar la relación entre hemisferios como una variable aleatoria que puede ser modelada y cuyos diferencias por individuo dependen exclusivamente de las características propias de este, una vez haya sido normalizado.

Otro aspecto importante y que es fundamental son las secuencias usadas para obtener las imágenes, de esto dependen los valores de los vóxeles obtenidos. En este estudio se usan dos secuencias, la FLAIR y la T1. Con la FLAIR se busca identificar hiperintensidades que puedan indicar algún tipo de anomalía y la secuencia T1 se uso para delimitar con mayor precisión el espacio descrito por cada cerebro para generar mapas de probabilidad con el método BIANCA (*Brain Intensity Abnormality Classification Algorithm*) descrito en la sección 2.4.

2.3. Coeficiente de correlación cruzada (NCC)

El coeficiente de correlación cruzada (NCC - *Normalized Cross Correlation*) ha sido ampliamente usado como una métrica para medir la similaridad entre imágenes [20]; esta medida aplica básicamente el mismo concepto del coeficiente de correlación de Pearson. Para este proyecto se busca crear un mapa de correlaciones entre hemisferios de manera que se compara vóxel a vóxel, donde las estimaciones de tendencia central y de dispersión son calculados a partir de los valores observados en cada hemisferio.

$$NCC = \frac{1}{N} \sum_i \frac{(x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y}$$

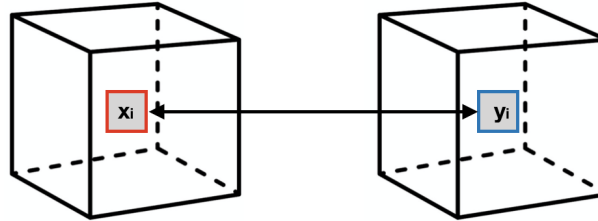


Figura 2.1: Coeficiente de correlación cruzada - NCC

Donde x_i, y_i son vóxeles de los volúmenes de ambos hemisferios y $\bar{x}, \bar{y}, \sigma_x, \sigma_y$ son sus respectivas medias y desviaciones estándar. El NCC puede interpretarse como la relación lineal entre la intensidad de vóxeles de dos volúmenes, en este caso, se trata de volúmenes que corresponden al hemisferio derecho e izquierdo.

Es posible obtener una versión robusta del NCC usando la mediana en lugar del promedio [13]. En este trabajo no se hace la prueba, pero en futuros proyectos puede resultar de interés calcular esta versión robusta del NCC para comparar resultados y costo computacional.

Con el NCC se estimaron mapas de correlación en el mismo espacio del cerebro para luego construir una imagen de cada hemisferio que puedan ser comparadas en magnitud-ubicación (vóxel-ubicación). Para hacer esto se siguieron los siguientes pasos:

1. Generación de un espacio común de comparación para ambos hemisferios con el cual puedan ser comparados luego cada uno por separado. En este caso el espacio en común es la correlación cruzada directa entre ambos hemisferios
2. Estimación del mapa de correlación del hemisferio derecho comparado con el espacio base obtenido en 1.
3. Estimación del mapa de correlación del hemisferio izquierdo comparado con el espacio base obtenido en 1.

Con esta medida se espera que hemisferios con una alta simetría tengan valores cercanos de los mapas de correlación. Es posible considerar el mapa NCC completo del cerebro al unir los mapas de cada hemisferio, de hecho, este ejercicio se realizó con el fin de averiguar si el espacio NCC completo permite detectar entre un cerebro sano y uno enfermo, después de aplicar métodos estadísticos.

2.4. BIANCA - *Brain Intensity Abnormality Classification Algorithm*

Lo que se ha mencionado hasta ahora solo considera trabajar con intensidades obtenidas de las neuroimágenes, sin embargo, resulta interesante observar cuáles serían los resultados si se usan mapas de probabilidad en lugar de las intensidades crudas o el mapa NCC. Con este enfoque se quiere evaluar si los mapas de probabilidad proporcionan datos más precisos al momento de comparar hemisferios de imágenes de sujetos sanos y patológicos.

BIANCA es un algoritmo supervisado que estima la probabilidad de que un vóxel tenga hiperintensidad sospechosa de enfermedad basado en sus k-vecinos más próximos. Se hace un entrenamiento con imágenes marcadas manualmente (se marcan las regiones patológicas de una neuroimagen corte por corte), para este entrenamiento se requieren entre 10 y 20 sujetos patológicos. BIANCA esta implementado en la biblioteca FSL [19]. En la figura 2.2 se puede observar un mapa de probabilidades después de aplicar el algoritmo a una imagen de un paciente diagnosticado con oligodendroglioma.

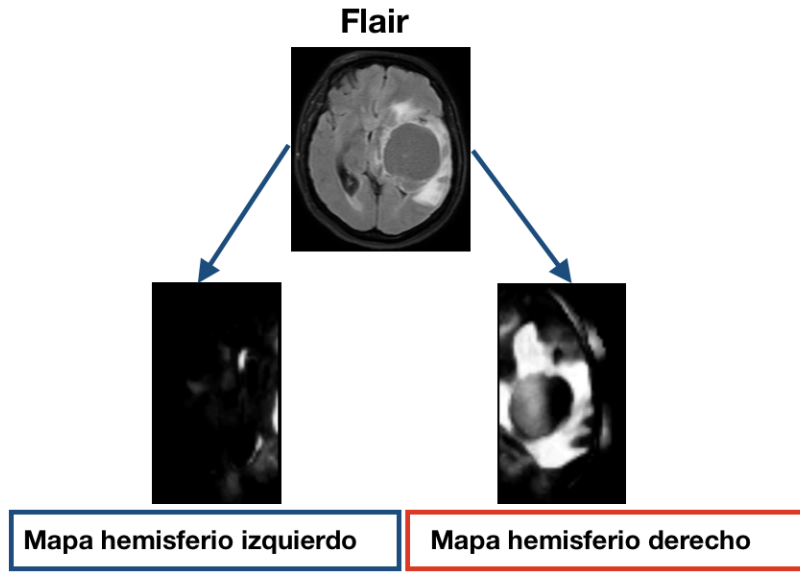


Figura 2.2: Mapa de probabilidades por hemisferio obtenidos con BIANCA, en un sujeto diagnosticado con oligodendroglioma

En este proyecto se hizo un entrenamiento con 20 imágenes marcadas por una médica radióloga con 9 años de experiencia en neuroimágenes. Observe que la marcación de una imagen en múltiples regiones a través de todos los cortes de una secuencia FLAIR equivale a tener cientos o miles de vóxeles marcados, es decir que con 20 neuroimágenes patológicas marcadas se tienen miles o millones de vóxeles para el entrenamiento.

2.5. Permutaciones de vóxeles entre hemisferios (HCP)

Se construyó una medida para comparar ambos hemisferios usando valores de vóxeles en el intervalo $[0, 1]$ y el mapa de probabilidades obtenido con BIANCA. En ambos casos se siguieron los siguientes pasos:

1. Sea n el número de vóxeles contenido en cada hemisferio, luego se genera una muestra aleatoria de números en el intervalo $[1, n]$ con reemplazo donde finalmente solo permanecen los valores únicos.
2. Selección en las imágenes de ambos hemisferios las posiciones generadas en el paso 1, después de haber colapsado las imágenes en un vector n – dimensional de $x * y * z$, donde x, y, z corresponde a las dimensiones originales de las imágenes de cada hemisferio.
3. Los vóxeles seleccionados para cada hemisferio en el paso 2 son intercambiados, obteniendo un par de vectores H_{right}, H_{left}

4. Cálculo de la norma L de los vectores obtenidos en 3 para calcular la razón $r(H_{right}, H_{left}) = L(H_{left}) / L(H_{right})$
5. Se Repite los pasos 1 a 5, k veces.

Se espera que $r(H_{right}, H_{left}) \approx 1$ bajo la hipótesis de que ambos hemisferios son altamente simétricos. Se hace un intercambio aproximadamente del 50 % de los datos entre hemisferios $k - veces$ evitando así que las diferencias que se observen en el test sean debidas al azar, es decir, es menos probable que valores de vóxeles de la misma magnitud ubicadas en diferentes regiones de los hemisferios den como resultado un valor del HCP cercanos a 1 indicando simetría entre hemisferios cuando en realidad no lo son. Este paso de intercambio entre valores de hemisferios que coincidan en las mismas posiciones le permite integrar a la prueba el concepto de ubicación espacial de manera implícita.

Aunque es relativamente sencillo el cálculo del test, su costo computacional es alto si se considera hacer para una muestra grande de imágenes, a pesar que el algoritmo podría ser escrito de manera que parte de los cálculos se hagan paralelamente [6].

2.6. Entropía de valores locales permutados entre hemisferios (HPLE)

En el HCP no se consideran efectos locales o de vecindario; con el HPLE se busca mejorar este aspecto al comparar los valores de intensidad de manera local por medio de los niveles de entropía después de haber permutado vóxeles locales entre hemisferios. En un cerebro sano se espera que de manera local la distribución de los valores observados de vóxeles en cada hemisferio tengan valores de entropía similares. Un aspecto importante de esta prueba son las permutaciones que se realizan previo al cálculo de la entropía; su importancia radica en que se le da mayor capacidad de encontrar diferencias a la prueba.

Considere que se seleccionan vóxeles en cada hemisferio y suponga que se quiere calcular el valor de la entropía sin hacer permutaciones como se muestra en la figura 2.3. Observe que la entropía calculada puede ser igual aunque la distribución de los valores de vóxeles sea espacialmente diferente. Ahora supongo que se permutan los valores y se vuelve a calcular la entropía, en este caso la incertidumbre daría valores diferentes (figura 2.4).

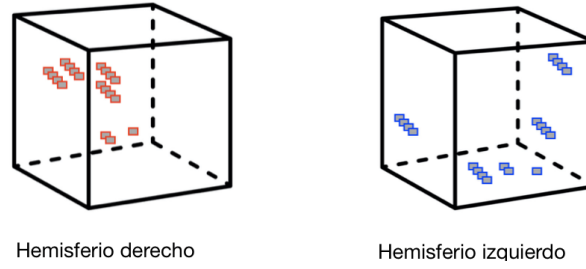


Figura 2.3: Selección de regiones en ambos hemisferios para calculo del HPLE, sin hacer permutaciones

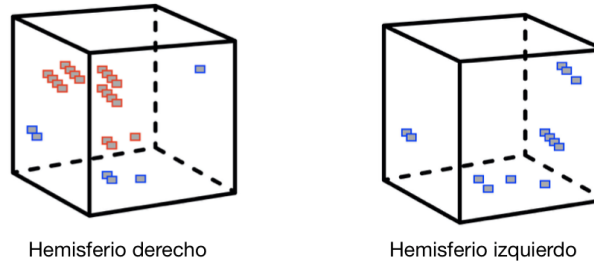


Figura 2.4: Selección de regiones en ambos hemisferios para calculo del HPLE, haciendo permutaciones

Para el cálculo de la entropía es necesario conocer las probabilidades asociadas a los valores observados; para la estimación de estas probabilidades se propone un procedimiento sencillo:

- Cálculo de percentiles en vóxeles seleccionados.
- Considere estos valores como valores referencia de su distribución.
- Repetir $100/n_{\text{percentiles}}$ cada valor de referencia.
- Finalmente, se calcula la probabilidad de cada valor referencia con los valores obtenidos en el punto anterior.

Observe que el cálculo de estas probabilidades a partir de un kernel es poco viable ya que muchos de los vóxeles tendrán valores iguales a cero en cada una de las regiones seleccionadas, además en cada región pueden haber pocos datos (en este caso se usaron regiones con 27 vóxeles). Por este motivo se propone este sencillo calculo, aunque obviamente se trata de un procedimiento experimental que podría ser mejorado.

En este trabajo se evitó incluir valores extremos como valores de referencia por lo que se usaron los percentiles 25, 50 y 75.

Después de describir la forma en que se calcularon las probabilidades, a continuación se listan los pasos seguidos para la estimación de la entropía:

1. Identificación de vóxeles con valores diferentes de cero en alguno de los dos hemisferios.
2. Construcción de cubos alrededor de los vóxeles identificados en el paso 1, en este caso se usaron cubos de $(3 \times 3 \times 3)$. A estos cubos se les llama subregiones.
3. Realizar las permutaciones de valores entre cada subregión (hay que recordar que se están comparando subregiones entre hemisferios que están en la misma ubicación espacial).
4. Calcular las probabilidades por hemisferio como se describió previamente.
5. Finalmente se estima la entropía propuesta por Shannon [47]

$$H(X) = - \sum_i p(x_i) \log_2 p(x_i)$$

Algo interesante del HPLE como indicador de similaridad es que usa la incertidumbre como medida de comparación entre hemisferios en regiones locales, lo que puede arrojar información adicional si se compara con el HCP. Además, los valores obtenidos en cada hemisferio pueden ser tratados como variables aleatorias con el fin de realizar pruebas estadísticas conocidas para detectar diferencias significativas.

2.7. Eigenfaces

Eigenfaces es un método que fue propuesto para reconocer imágenes de rostros humanos; es en esencia una variación del método de componentes principales (PCA por sus siglas en inglés) con el que se busca encontrar un subespacio que resuma de la mejor manera el espacio generado por un conjunto de imágenes [48, 57]. Por esta razón, para entender un poco mejor el método de Eigenfaces se hace primero una descripción del PCA.

2.7.1. Análisis de componentes principales (PCA)

El PCA está dentro de los métodos no supervisados y es uno de los más usados para la reducción de dimensionalidad, su filosofía es tratar de reducir

complejidad (dimensionalidad) conservando información estadística del conjunto de datos (variabilidad). En términos generales se hace la pregunta ¿cuáles son las direcciones de mayor variabilidad en el conjunto de datos?

En la práctica se parte de un conjunto de variables $X = [x_1, x_2, \dots, x_n]$ correlacionadas que se quieren transformar en un nuevo conjunto de variables incorrelacionadas (componentes) que son combinaciones lineales de las variables originales; se buscan combinaciones lineales de las columnas de X con máxima varianza, esto es $\text{var}(Xa) = a^T S a$, donde S es la matriz de covarianzas muestral y a es un vector de constantes; por tanto, encontrar la combinación lineal con máxima varianza equivale a resolver $a^T S a$.

Una forma de resolver esta ecuación cuadrática es mediante el método de multiplicadores de Lagrange imponiendo restricciones de modularidad sobre a y haciendo que las covarianzas entre componentes sea igual a cero. Otra forma equivalente de hacerlo es mediante la descomposición en valores singulares o descomposición espectral de la matriz $(n - 1)S = X_c^T X_c$ donde X_c es la matriz X centralizada [17, 28]. En este trabajo se usó la descomposición espectral.

Existen otras técnicas para la reducción de dimensionalidad que en futuros trabajos se podrían aplicar al problema tratado en este proyecto. El algoritmo *t-distributed stochastic neighbor embedding* (t-SNE) puede ser una opción interesante ya que es posible encontrar combinaciones que no necesariamente son lineales como lo plantea el PCA [9]

2.7.2. Eigenfaces: PCA modificado

Uno de los problemas de usar PCA directamente sobre las imágenes es que se trabaja sobre la matriz de covarianzas generada por estas, la cual puede alcanzar dimensiones bastante altas. Considere una imagen en $2 - d$ de tamaño $N \times N$, por ejemplo, 250×250 píxeles, luego cada imagen está definida en un espacio con una dimensión de 62500, por lo que la manipulación de la matriz de covarianzas requerida para los análisis de PCA se hace computacionalmente inviable con dimensiones de 62500×62500 .

Con Eigenfaces se propone extraer información relevante de un conjunto de imágenes, pero en este caso se hace a partir de las M imágenes que se tengan y no del número de dimensiones descrito por la imagen. Este es un salto importante ya que reduce los cálculos necesarios en el caso que $M < N$ (que es en la mayoría de los casos). A continuación se describen de manera breve los principios y pasos a seguir para obtener los subespacios mencionados [58]:

1. Calcular la imagen media o mediana de las neuroimágenes, Ψ .
2. Representar cada imagen en un vector $n - dimensional$, Γ (colapsar las imágenes).
3. Centralizar cada imagen $\Phi_i = \Gamma_i - \Psi, i = 1, 2, \dots, M$.
4. Sea $A = [\Phi_1, \Phi_2, \dots, \Phi_M]$, con M : número de imágenes en el conjunto de datos. La matriz de covarianzas se define como $C = A^T A$
5. Encontrar los valores y vectores propios asociados a C , tal que $A^T A v_i = w_i v_i$ (descomposición en valores singulares), con v_i : vector propio de C y w_i : valor propio asociado.
6. Seleccionar los vectores propios asociados a los k valores propios más altos.

Puede decirse que las imágenes que forman el subespacio son las características más representativas construidas a partir de las imágenes originales con las que se puede hacer una reconstrucción con un alto grado de precisión; entendiendo la precisión como el grado de variabilidad explicado por el subespacio (es el mismo concepto usado en PCA) [57]. Igual que en el PCA, se realiza una descomposición en valores singulares para encontrar el mejor subespacio, los que reciben el nombre de eigenfaces y eigenvalues. Los eigenvalues indican el grado de variabilidad capturado por cada eigenfaces, donde cada eigenfaces corresponde a una imagen que es combinación lineal de las imágenes originales. Observe que el número posible de vectores propios es igual al número de imágenes en la muestra.

En el presente trabajo se usó el método de Eigenfaces desde dos perspectivas para tratar de evaluar en que casos es más apropiado usarlo para determinar si una imagen proviene de un cerebro sano o enfermo:

- A Se estimó un subespacio en cada hemisferio para sanos y enfermos, luego se calcularon las distancias de las proyecciones de cada hemisferio a su respectivo espacio propio y finalmente se estiman las diferencias de las distancias. Bajo este enfoque se espera que las diferencias de las distancias de los sanos sean menores que la de los enfermos. Aquí se usaron los valores de intensidades obtenidos con la secuencia FLAIR, aunque también es posible hacerlo con los mapas de probabilidad obtenidos con BIANCA.
- B Se usó el mapa de correlaciones generado con el NCC, observe que las dimensiones del NCC es igual al de la neuroimagen original, pero esta fue construida por hemisferios (ver sección 2.3). Con estos valores se

construyó un subespacio para las imágenes de pacientes sanos, luego se proyectan ambos conjuntos de datos al espacio propio, después con estas proyecciones se construyó un modelo de clasificación. Este mismo ejercicio se hizo con el mapa de probabilidades.

Como se mencionó, el "secreto" de Eigenfaces es que trabaja con la matriz de covarianzas definida por $C_{Eigenfaces} = A^T A$ y no con $C_{PCA} = A A^T$, pero ¿cómo se relacionan los espacios propios obtenidos con cada una de las dos matrices? veamos:

Suponga que se encuentran los vectores propios u_i y valores propios w_i para C_{PCA} tal que $A A^T u_i = w_i u_i$, luego

$$\begin{aligned} A^T A v_i &= w_i v_i \\ A A^T A v_i &= w_i A v_i \\ C_{PCA} A v_i &= w_i A v_i \\ C_{PCA} u_i &= w_i u_i \end{aligned} \tag{2.1}$$

donde $u_i = A v_i$, de la ecuación 2.1 se observa que los valores propios de $C_{Eigenfaces}$ y C_{PCA} son iguales, y los vectores propios se relacionan mediante $u_i = A v_i$ [58]

Una de las preguntas que surge a partir del uso de Eigenfaces es ¿qué pasa si el espacio propio se construye a partir de imágenes diferentes, que tanto afecta los resultados de los modelos finales? esto es un aspecto crucial que se abordó en este trabajo. Para este fin se construyeron múltiples subespacios de sujetos sanos con muestra de imágenes diferentes. Los resultados de los modelos obtenidos con estos subespacios pueden ser vistos como un ejercicio de reproducibilidad que da respuesta a la pregunta expuesta.

En la literatura se han propuesto otros métodos lineales y no lineales para la reducción de dimensionalidad en imágenes [4], sin embargo, en este estudio solo se usó el método PCA modificado.

2.8. Modelos de clasificación supervisados

Los modelos de clasificación supervisados han sido usados en problemas de predicción en diversos campos de estudio [32]. Con la introducción de nuevos y mejorados algoritmos de optimización y más capacidad de computo, se

tienen disponibles varios tipos de modelos en paquetes como R y Python. Entre ellos se encuentra el modelo logístico, árboles de decisión, *random forest*, máquinas de soporte vectorial, Naive Bayes, redes neuronales, entre otros [24].

La filosofía detrás de los modelos de clasificación supervisados es que al tener un conjunto de datos etiquetados con la clasificación verdadera, se trata de construir un modelo que sea capaz de detectar patrones en los datos que le permitan distinguir entre los grupos definidos por las etiquetas. Para la detección de estos patrones o características, en cada tipo de modelo en particular se proponen funciones que deben ser optimizadas para estimar los parámetros definidos[26].

Lo que se conoce hoy como *machine learning*, se podría decir que es una extensión de los modelos estadísticos conocidos, pero con variaciones y mejoras en los algoritmos para resolver las funciones de optimización, por ejemplo, el algoritmo de gradiente descendiente [49]. Otro de los grandes hitos en *machine learning* unido a la capacidad computacional es el desarrollo de redes neuronales como herramienta de predicción, consiguiendo construir topologías bastante complejas con millones de parámetros que permiten resolver problemas que antes eran intratables [1].

Es claro que la era del *machine learning* es la era de las predicciones, pero poco se ha desarrollado para resolver problemas de causalidad (cuáles son los factores causales y cuales son sus pesos o importancia), por lo que aun queda mucho por trabajar en esta línea, además de mejorar los resultados obtenidos hasta ahora con los modelos no supervisados.

Se usaron los siguientes modelos de clasificación para tratar de predecir si una neuroimagen proviene de un cerebro sano o uno patológico: Logístico, análisis discriminante, árboles de decisión, *random forest*, máquinas de soporte vectorial, *Adaboost classifier*, *Stochastic boosting classifier* y Naive Bayes.

Para la regresión logística se usó el algoritmo Newton-CG (*Newton-conjugate gradient*) [44] implementado en la biblioteca Sklearn de python, para el modelo *random forest* se usaron 100 árboles. Los demás parámetros usados en los modelos son los que tienen por defecto en la biblioteca Sklearn.

Para evaluar la capacidad predictiva se usaron cinco indicadores:

- **Accuray:** Medida global de lo correctamente predicho, se puede presentar como proporción o porcentaje.

- **Precisión:** Es la razón entre los casos positivos predichos correctamente y el total de casos predichos como positivos.
- **Recall:** Razón entre los casos positivos predichos correctamente y el total de casos positivos en la muestra.
- **F1-score:** Es un promedio ponderado del *recall* y la precisión.
- **Área bajo la curva roc:** Indicador que mide la capacidad predictiva de un modelo.

Capítulo 3

Objetivos

3.1. General

Clasificar neuroimágenes de resonancia magnética en las categorías sano o patológico mediante métodos estadísticos que puedan ser implementados en un ambiente de producción.

3.2. Específicos

- Revisar y seleccionar imágenes que no tengan artefactos y que no corresponda a pacientes que fueron sometidos a cirugías de neurología.
- Establecer flujo de trabajo para la extracción y procesamiento de imágenes.
- Aplicar algoritmo BIANCA para obtener mapa de probabilidades.
- Ajustar modelos de clasificación, calcular el HCP y HPLE con los datos obtenidos del NCC, Egenfaces, BIANCA y la secuencia FLAIR.

Capítulo 4

Métodos

4.1. Tipo de estudio

Este es un estudio de corte transversal.

4.2. Población

La población son las imágenes de resonancia magnética de la especialidad de neurología que han sido adquiridas en Servicios de Salud - San Vicente Fundación entre los años 2015 y 2019. Se cuenta con aproximadamente 15000 neuroimágenes.

4.3. Muestreo

Para el cálculo del HCP y HPLE se tomó un conjunto de imágenes para las que se tenía disponible una clasificación hecha por radiólogos expertos en neurología, por lo que se trata de un muestreo por conveniencia.

Para el calculo del NCC, se uso un conjunto de imágenes que contaban con una secuencia FLAIR axial, y para BIANCA con conjunto que contaban con las secuencias FLAIR axial y T1.

4.4. Tamaño de la muestra

En total se contaba con 750 neuroimágenes, de las cuales 381 fueron usadas para la construcción de espacios propios de sujetos sanos con la metodología Eigenfaces, y 369 (212 sanos, 157 patológicas) para ajustar modelos y hacer cálculos del HCP y HPLE.

4.5. Selección de la muestra

La selección de la muestra se hizo de manera aleatoria, aunque el tamaño de la muestra es no probabilística.

4.6. Criterios de inclusión

Se incluyeron estudios de neuroimágenes que cuentan con la secuencia FLAIR. Para aplicar BIANCA se seleccionaron los estudios que tenían las secuencias T1 y FLAIR.

4.7. Criterios de exclusión

Se excluyeron las imágenes de sujetos que se les habían hecho alguna intervención quirúrgicas en el cerebro, tampoco se consideran aquellas imágenes con artefactos [46].

4.8. Fuentes de información

Se usó una fuente de datos secundaria que consiste en imágenes DICOM alojadas en un PACS de Servicios de Salud – San Vicente Fundación.

4.9. Control de sesgos y overfitting

Para evitar sesgos de selección se incluyen estudios con diferentes tipos de tumores, enfermedad cerebrovascular y enfermedad de sustancia blanca, buscando que el modelo de predicción aprenda a detectar anomalías que provengan de diferentes patrones. Para el ajuste de modelos de clasificación se incluyó un 42 % de sujetos enfermos en la muestra para evitar desbalanceo en los datos.

Para evitar el overfitting, la construcción de los espacios propios de pacientes sanos usados para ajustar modelos de clasificación con el enfoque B descrito en 2.7.2, se hizo con un conjunto de imágenes que luego no es usado para el ajuste de los modelos. Es decir, que los datos usados en los modelo de clasificación no provienen de las imágenes usados en la construcción del espacio propio.

4.10. Limitaciones del estudio

Este estudio solo contempla la clasificación en dos categorías (sano o patológicas), no se trabaja con escalas ordinales o continuas para la priorización, lo cual puede ser importante abordarlo en otros estudios.

Tampoco se cuenta con un tamaño de muestra estimado bajo principios probabilísticos o del criterios de la complejidad óptima de la muestra del aprendizaje PAC (*Probably approximately correct learning*) [22]. Para futuros estudios se recomienda aumentar el tamaño de la muestra y evaluar de nuevo los métodos propuestos.

4.11. Procesamiento de imágenes

Las neuroimágenes fueron tomadas en resonadores Philips Ingenia de 1.5 y 3.0 Tesla, el procesamiento o preparación de las imágenes fueron realizadas con la biblioteca FSL. El procesamiento de las imágenes consiste en la extracción del cerebro (excluyendo tejido que no es parte del cerebro y hueso), una corrección por movimientos y una normalización al espacio estándar MNI 152 [50, 8]. Adicionalmente se extrajo cada hemisferio del cerebro (ver figura 4.1).

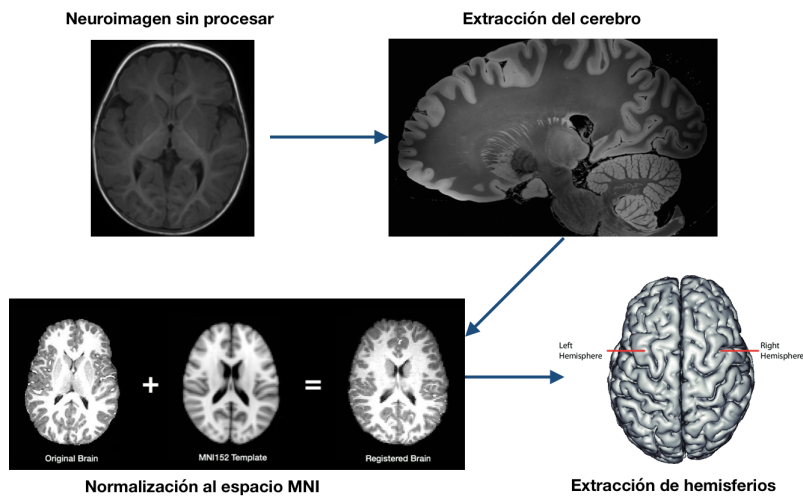


Figura 4.1: Procesamiento de neuroimágenes de RM.

4.12. Preparación y partición de datos para el ajuste de modelos

Para ajustar los modelos de clasificación bajo el enfoque **B** descrito en la sección 2.7.2, se construyeron 200 espacios propios, después de seleccionar de manera aleatoria en cada ocasión un 60 % de las 381 imágenes destinadas para este ejercicio. Luego de obtener las proyecciones al espacio propio de las 369 imágenes usadas para los modelos de clasificación, se dividieron los datos en dos conjuntos, uno para el entrenamiento (70 %) y otro de prueba (30 %), además, se estandarizaron las variables que se ingresaron a los modelos.

Observe que el ajuste de modelos a partir de múltiples espacios propios contruidos con imágenes distintas, permite obtener resultados que son reproducibles, es decir, que con otros conjuntos de datos usados por investigadores que deseen replicar el ejerció hecho en este proyecto los resultados no deberían variar demasiado.

Para el caso del ajuste de modelos bajo el enfoque **B** usando los mapas de probabilidad obtenidos con BIANCA, se usaron 271 imágenes (196 sanos y 75 patológicos), se siguió el mismo principio que se acaba de describir (selección del 60 % para construir múltiples subespacios, y partición en 30 % y 70 % para prueba y entrenamiento respectivamente).

4.13. Variables en el estudio

Cuadro 4.1: Variables en el estudio

Variable	Nivel de medición	Unidades
Grupo	sano/patológico	NA
Hemisferio derecho	Escala	mm^3
Hemisferio izquierdo	Escala	mm^3
BIANCA	Escala	mm^3
NCC	Escala	mm^3
HCP	Escala	NA
HPLE	Escala	NA

4.14. Consideraciones éticas

El proyecto considera las normas internacionales y nacionales para la investigación con seres humanos, entre ellos el Código del Tribunal Internacional

de Núremberg[40], en el que se exponen los principios básicos para investigación en seres humanos tales como: consentimiento informado, mínimos riesgos, máximos y derecho al retiro voluntario. El informe Belmont [7] que considera los principios de respeto, beneficencia y justicia. La declaración de Helsinki de la Asociación Médica Mundial la cual afirma que el bienestar de la persona prevalece sobre otros intereses y rescata el deber de proteger la vida, la salud, la dignidad, la integridad, el derecho a la autodeterminación, la intimidad y la confidencialidad de la información. De acuerdo con la resolución número 8430 de 1993 de la República de Colombia expedida por el Ministerio de Salud y tal como dicta en el Título II, capítulo 1, artículo 11, esta es una investigación con riesgo mínimo ya que se emplea el registro de datos existentes (fuente de dato secundaria) anonimizados.

Nota: Este estudio se centra en realizar procesamiento de imágenes, aplicar métodos estadísticos y matemáticos en relación con el manejo de los datos obtenidos, para los cuales se siguen respetando los criterios de confidencialidad y de uso científico y académico de la información.

Este proyecto fue aprobado por el Comité de Ética de la Fundación Hospitalaria San Vicente de Paúl mediante el acta número 19-2020.

Capítulo 5

Resultados

En primer lugar se muestran los resultados obtenidos con el test de permutaciones entre hemisferios (HCP), luego se presentan lo observado con los valores de entropía de valores locales permutados entre hemisferios (HPLE) y finalmente son presentados los resultados obtenidos con Eigenfaces.

Antes de entrar en detalles, se realiza una exploración inicial de las diferencias en valor absoluto de los hemisferios para cada grupo por separado, en sanos y patológicos. Este primer acercamiento permite dar una idea global de las distribuciones de las diferencias entre hemisferios en cada grupo. Para cada imagen se repitió 1000 veces el siguiente ejercicio:

1. Se generó una muestra aleatoria de n valores con reemplazo, donde solo permanecen los valores únicos, los cuales corresponden a índices de las posiciones de valores en los vectores de intensidades de las imágenes colapsadas.
2. Los índices generados en 1 se seleccionaron en cada hemisferio y se calculó el valor absoluto de sus diferencias $D(H_r, H_l)$ (D es un vector de tamaño k , donde k corresponde al tamaño del vector generado en 1)
3. Se calculó la proporción de valores en D que son mayores a cero

$$\hat{p} = \frac{1}{k} \sum_i I(D_i); i = 1, 2, \dots, k \quad (5.1)$$

donde

$$I(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{si } x = 0 \end{cases}$$

4. Se calculó el promedio de los valores mayores a cero en D

$$\bar{d} = \frac{1}{l} \sum_j D_j; j = 1, 2, 3, \dots, l$$

donde l es la cantidad de elementos tal que $D > 0$

5. Se estimó $\bar{d}p = \bar{d} + \hat{p}$

Los resultados de estas estimaciones se resumen en la figura 5.1

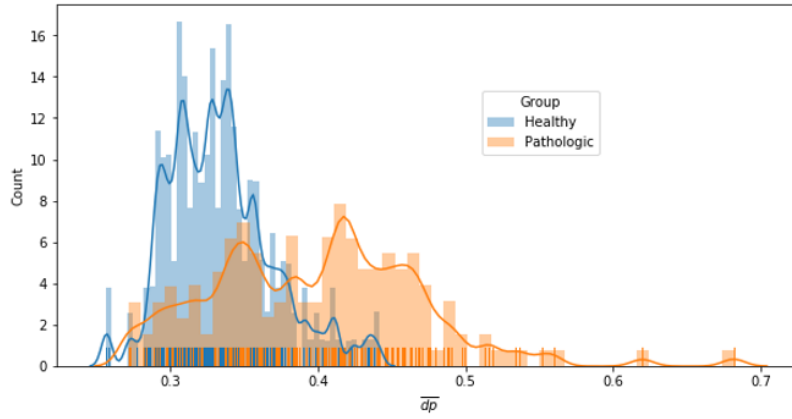


Figura 5.1: Distribución de las diferencias entre hemisferios ajustadas, según el grupo.

Se observa que los valores de $\bar{d}p$ en el caso de los patológicos tiene una distribución con un rango más amplio con una cola a la derecha (5.1). Al definir el umbral $tr = \max(\bar{d}p_{healthy})$, se encontró que el 26 % de los valores de $\bar{d}p_{pathologic}$ superaron este umbral, es decir que con esta medida se logra diferenciar solo un 26 % de los patológicos, el 74 % restante se confunde con los valores de los sanos. Este es un resultado importante porque sugiere que se deben encontrar características en los datos que permitan hacer una buena “separación” de ambos grupos, sabiendo que el calculo de una diferencia directa no es suficiente, aunque sea ajustada y/o escalada (al escalar $\bar{d}p$ se obtuvieron resultados similares).

5.1. HCP

El HCP se calculó con valores de intensidades normalizados en el intervalo $[0, 1]$ y con el mapa de probabilidades obtenido con BIANCA. A continuación se presentan los dos resultados.

5.1.1. HCP con valores de intensidades normalizados

Para cada imagen se obtuvieron 500 valores del test $r(H_R, H_L) = L(H_R)/L(H_L)$ como se menciona en la sección 2.5, usando la norma L_1 .

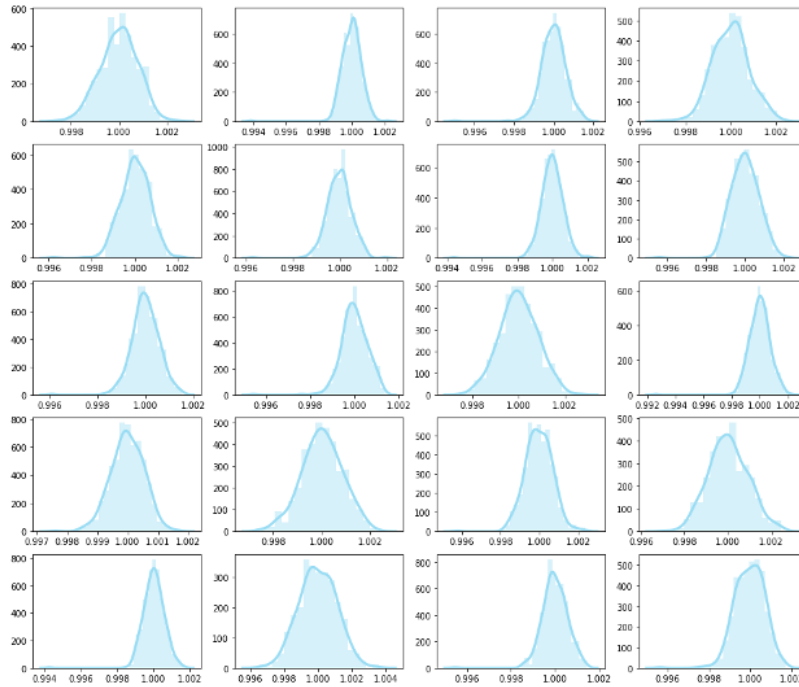


Figura 5.2: Distribución del test HCP en 20 neuroimágenes de sujetos sanos.

En el caso de los sujetos sanos la distribución observada para el test es coherente con lo que se planteó en la hipótesis (valores que están alrededor de 1). Sin embargo, para los patológicos se esperaba que la distribución no estuviera centrada en 1 (figuras 5.2 y 5.3). Un aspecto que se debe resaltar es que aunque en efecto para ambos tipos de imágenes el test HCP está alrededor de 1, para los patológicos se observa un rango de valores más amplia y una varianza mayor comparado con los sanos (tabla 5.1)

Cuadro 5.1: Medidas de resumen del test HCP de sanos y patológicos.

Medida de resumen	Sanos	Patológicos	Patológicos/sanos
Media	1.0000	1.0	1.0000
Mediana	1.0005	1.0006	1.0000
Desviación estándar	0.0021	0.0169	8.0476
Rango	0.0062	0.0142	2.2903
Rango intercuartílico	0.0009	0.0014	1.5555
Mínimo	0.996	0.9923	0.9963
Máximo	10.022	10.066	1.0044

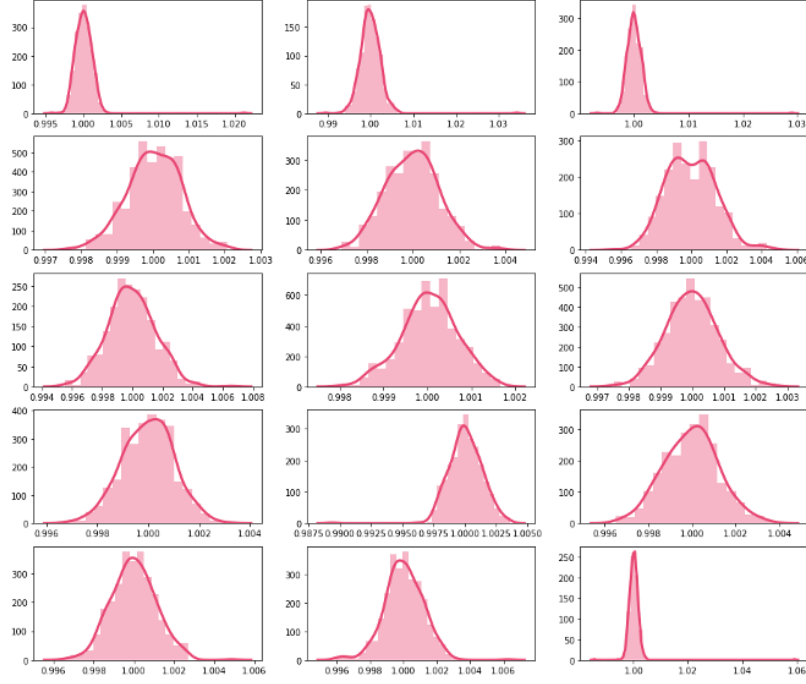


Figura 5.3: Distribución del test HCP en 20 neuroimágenes de sujetos patológicos.

De la tabla 5.1 se observa que la desviación estándar de los patológicos es 8 veces la de los sanos, su rango 2 veces mayor y el rango intercuartílico 1.5 veces. Dado este comportamiento se podría tratar de plantear una prueba que se base en la distribución del mínimo, el máximo y la desviación; como ejercicio se establecieron valores de referencia de los sujetos sanos. El valor del percentil 90 del mínimo de HCP (min), el percentil 10 del máximo (max) y el percentil 10 de la desviación (min_{std}), luego se establecieron las siguientes reglas de clasificación:

$$\hat{p}_{min} = \frac{1}{n} \sum_i I_{min}(HCP); i = 1, 2, \dots, n \quad (5.2)$$

$$\hat{p}_{max} = \frac{1}{n} \sum_i I_{max}(HCP); i = 1, 2, \dots, n \quad (5.3)$$

donde

$$I(x)_{min} = \begin{cases} 1 & \text{si } x \leq min \\ 0 & \text{si } x > 0 \end{cases}$$

$$I(x)_{max} = \begin{cases} 1 & \text{si } x \geq \min \\ 0 & \text{si } x < 0 \end{cases}$$

$$Zonas(\hat{p}_{min}, \hat{p}_{max}, \hat{\sigma}_{HCP}) = \begin{cases} \text{patológica} & (\hat{p}_{min} > 0,01 \vee \hat{p}_{max}) > 0,01 \wedge (\hat{\sigma}_{HCP} > \min_{std}) \\ \text{indecisión} & (\hat{p}_{min} > 0,01 \vee \hat{p}_{max}) > 0,01 \wedge (\hat{\sigma}_{HCP} \leq \min_{std}) \\ \text{sanos} & \text{otro caso} \end{cases} \quad (5.4)$$

De esta manera se clasificaron los estudios como sigue (tabla 5.2):

Cuadro 5.2: Clasificación según la distribución del test HCP

Grupo	Zonas de decisión		
	Sanos	Patológica	Indecisión
Sanos	149 (70 %)	22 (11 %)	41 (19 %)
Patológicos	34 (22 %)	96 (61 %)	27 (17 %)

De la tabla 5.2 se observa que la clasificación correcta de patológicos y sanos es baja (61 % y 70 % respectivamente). Si se asume que los estudios que están en la zona de indecisión son patológicos por sospecha, el *accuracy* sería igual al 73.7 %, resultado que sigue siendo poco satisfactorio para usarlo en la practica clínica.

Los resultados obtenidos con el HCP evidencian que esta prueba tal como fue propuesta tiene dificultades para separar de manera satisfactoria ambos grupos, aunque se logran resultados mucho mejores comparado con las diferencias ajustadas que se presentaron al principio del capitulo. Es posible que se deban agregar otras características a la prueba en busca de capturar más diferencias.

Se realizó el mismo ejercicio con las intensidades mayores que cero, el comportamiento fue similar (no se presentan los resultados).

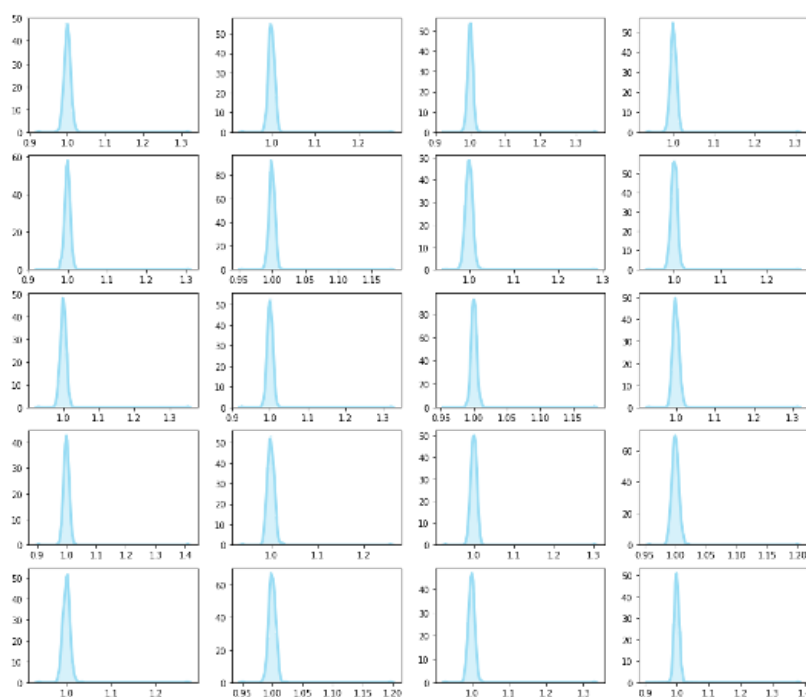
5.1.2. HCP con mapas de probabilidad

En este caso también se observan valores que oscilan alrededor de 1 para ambos grupos (figuras 5.4 y 5.5), sin embargo, no se observan tantas diferencias en el rango de valores del HCP, pero si en la varianza (tabla 5.3). Por tanto la matriz de confusión se construyó teniendo en cuenta solo la desviación como parámetro (tabla 5.4).

Cuadro 5.3: Medidas de resumen del test HCP de sanos y patológicos - BIANCA.

Medida de resumen	Sanos	Patológicos	Patológicos/sanos
Media	1.0005	10.005	1.0000
Mediana	1.0048	10.048	1.0000
Desviación estándar	0.0804	0.1451	1.8047
Rango	0.3524	0.3339	0.9475
Rango intercuartílico	0.0096	0.0095	0.9896
Mínimo	0.9361	0.9352	0.9990
Máximo	1.2885	1.2691	0.9849

De la tabla 5.3 se observa que la desviación es la única variable en la que el HCP de los patológicos muestra diferencias notables, siendo este 1.8 veces la desviación de los sanos.

**Figura 5.4:** Distribución del test HCP en 20 neuroimágenes de sujetos sanos - BIANCA.**Cuadro 5.4:** Clasificación según la distribución del test HCP

Grupo	Zonas de decisión	
	Sanos	Patológica
Sanos	20 (10 %)	176 (90 %)
Patológicos	16 (33 %)	59 (77 %)

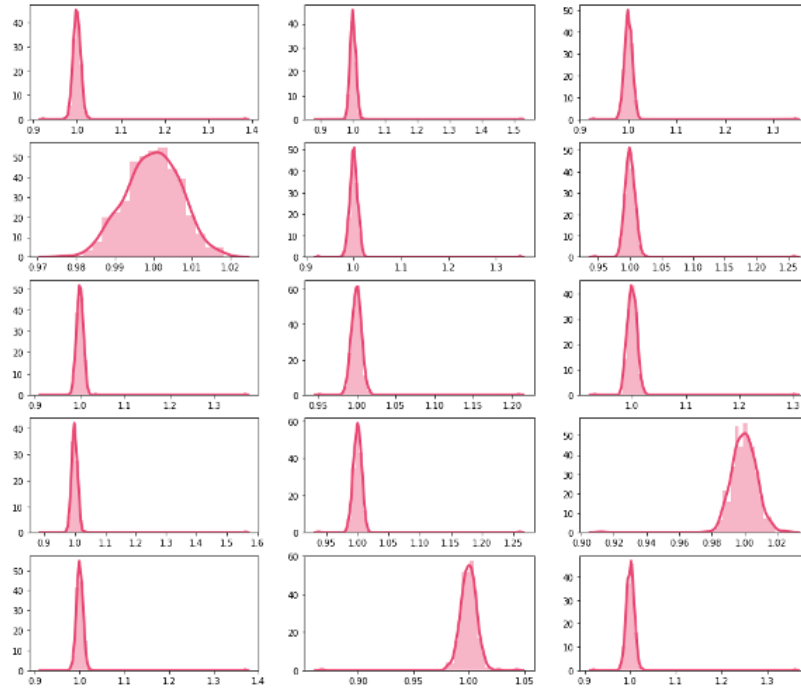


Figura 5.5: Distribución del test HCP en 20 neuroimágenes de sujetos patológicos - BIANCA.

La matriz de clasificación se construyó en este caso con la desviación estándar como único parámetro para dicha clasificación. En este caso los resultados fueron deficientes en la detección de sanos, para los patológicos se observan mejores resultados (77 % de casos detectados) (tabla 5.4). Estos valores sugieren que el mapa de probabilidades obtenido con BIANCA no es una buena opción para calcular el test HCP con un *accuracy* del 21.3 %.

5.2. HPLE

El test HPLE fue calculado con los valores de intensidades normalizados en el intervalo $[0, 1]$. En la figura 5.6 se muestra la distribución de las diferencias entre hemisferios, se observa que de nuevo las neuroimágenes patológicas son las que mayor variabilidad presentan.

A diferencia del HCP, en el HPLE se tienen valores para ambos hemisferios, por lo que es posible aplicar pruebas estadísticas paramétricas y no paramétricas para evaluar si hay diferencias significativas entre hemisferios en cada imagen. Para evaluar si hay diferencias se aplicó el estadístico *t* para varianzas diferentes y el test *U* de Mann-Whitney, con el primero se busca comparar las

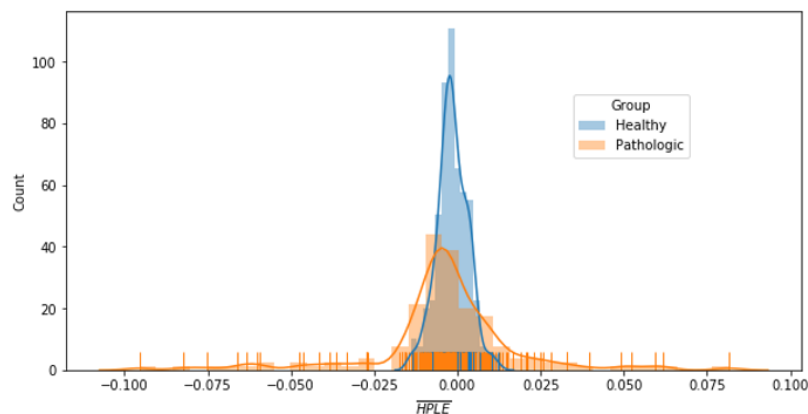


Figura 5.6: Distribución del HPLE entre hemisferios, según el grupo.

medias del HPLE de ambos hemisferios y con el segundo se trata de averiguar si ambas muestras provienen de la misma distribución [11].

Con ambas pruebas se obtuvo una capacidad predictiva similar del HPLE (tablas 5.5 y 5.6). En general se puede decir que el HPLE tiene un bajo rendimiento para diferentes niveles de significancia (figuras 5.7 y 5.8).

En ambos casos se observó que a medida que aumenta el nivel de significancia, también lo hace el *Recall*, pero disminuyen los demás indicadores (figuras 5.7 y 5.8).

En general, al tomar un nivel de significancia del 5 %, el *accuracy* no superó el 66 %.

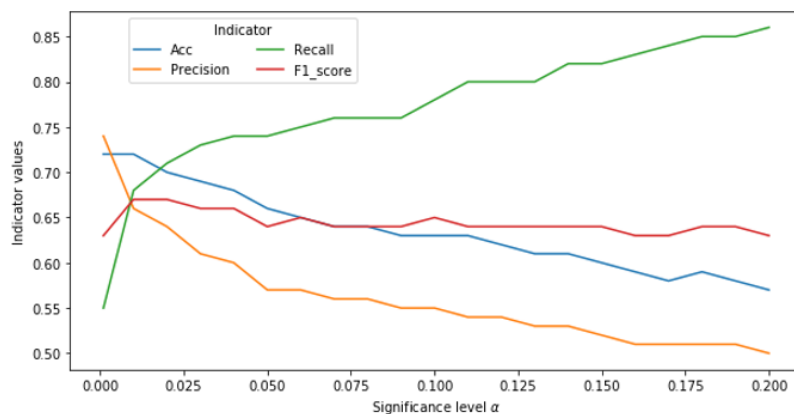


Figura 5.7: Tendencia de indicadores de desempeño del HPLE según el nivel de significancia, aplicando prueba U de Mann-Whitney.

Cuadro 5.5: Capacidad predictiva del HPLE según niveles de significancia, aplicando prueba U de Mann-Whitney.

α	VP	VN	FP	FN	Accuraccy	Precision	Recall	F1-score
0.001	86	181	31	71	0.72	0.74	0.55	0.63
0.01	106	158	54	51	0.72	0.66	0.68	0.67
0.02	112	148	64	45	0.70	0.64	0.71	0.67
0.03	115	138	74	42	0.69	0.61	0.73	0.66
0.04	116	135	77	41	0.68	0.60	0.74	0.66
0.05	116	126	86	41	0.66	0.57	0.74	0.64
0.06	117	122	90	40	0.65	0.57	0.75	0.65
0.07	119	119	93	38	0.64	0.56	0.76	0.64
0.08	120	117	95	37	0.64	0.56	0.76	0.64
0.09	120	112	100	37	0.63	0.55	0.76	0.64
0.10	123	110	102	34	0.63	0.55	0.78	0.65
0.11	125	107	105	32	0.63	0.54	0.80	0.64
0.12	125	104	108	32	0.62	0.54	0.80	0.64
0.13	126	100	112	31	0.61	0.53	0.80	0.64
0.14	129	96	116	28	0.61	0.53	0.82	0.64
0.15	129	91	121	28	0.60	0.52	0.82	0.64
0.16	130	89	123	27	0.59	0.51	0.83	0.63
0.17	132	83	129	25	0.58	0.51	0.84	0.63
0.18	134	83	129	23	0.59	0.51	0.85	0.64
0.19	134	81	131	23	0.58	0.51	0.85	0.64
0.20	135	76	136	22	0.57	0.5	0.86	0.63

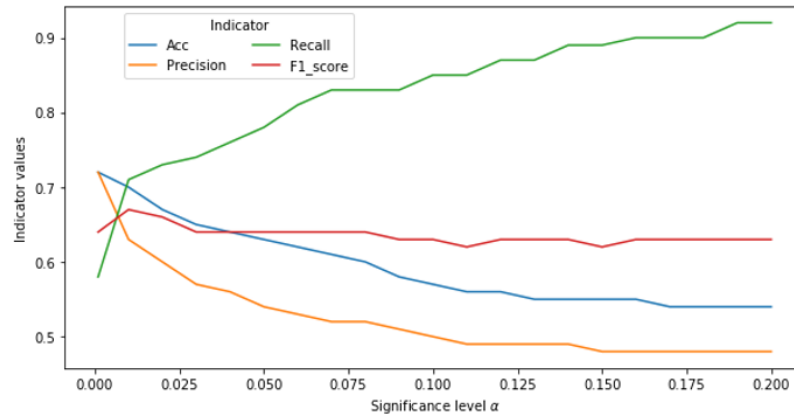


Figura 5.8: Tendencia de indicadores de desempeño del HPLE según el nivel de significancia, aplicando prueba t de student.

5.3. Eigenfaces

Como se menciono, se realizaron dos tipos de análisis con Eigenfaces, en el primero es construyendo subespacios por hemisferios y grupo, en el segundo

Cuadro 5.6: Capacidad predictiva del HPLE según niveles de significancia, aplicando la prueba t de student.

α	VP	VN	FP	FN	Accuraccy	Precision	Recall	F1-score
0.001	91	176	36	66	0.72	0.72	0.58	0.64
0.01	111	148	64	46	0.70	0.63	0.71	0.67
0.02	115	134	78	42	0.67	0.60	0.73	0.66
0.03	116	124	88	41	0.65	0.57	0.74	0.64
0.04	120	117	95	37	0.64	0.56	0.76	0.64
0.05	123	109	103	34	0.63	0.54	0.78	0.64
0.06	127	101	111	30	0.62	0.53	0.81	0.64
0.07	130	94	118	27	0.61	0.52	0.83	0.64
0.08	131	89	123	26	0.60	0.52	0.83	0.64
0.09	131	84	128	26	0.58	0.51	0.83	0.63
0.1	133	77	135	24	0.57	0.50	0.85	0.63
0.11	134	72	140	23	0.56	0.49	0.85	0.62
0.12	136	70	142	21	0.56	0.49	0.87	0.63
0.13	137	67	145	20	0.55	0.49	0.87	0.63
0.14	139	65	147	18	0.55	0.49	0.89	0.63
0.15	140	62	150	17	0.55	0.48	0.89	0.62
0.16	141	61	151	16	0.55	0.48	0.90	0.63
0.17	141	59	153	16	0.54	0.48	0.90	0.63
0.18	142	57	155	15	0.54	0.48	0.90	0.63
0.19	145	55	157	12	0.54	0.48	0.92	0.63
0.2	145	53	159	12	0.54	0.48	0.92	0.63

es construyendo un subespacio según el grupo pero sobre todo el cerebro.

Para cada caso la selección de los Eigenfaces o vectores propios se hizo reteniendo los eigenvalues de mayor valor que retuvieran el 99 % de la variabilidad. El valor de 99 % se elige ya que el interés principal no es la reducción de dimensionalidad, sino tratar de encontrar un espacio propio capaz de capturar la información más relevante, para luego hacer las proyecciones de cualquier neuroimagen sobre este, buscando encontrar que tanta información relevante (con respecto al subespacio) contiene la imagen proyectada.

Para la construcción del espacio propio de los sujetos sanos con todo el cerebro a partir de los valores de intensidad de la secuencia FLAIR axial, se separó un conjunto de 381 imágenes que solo se usaron para este fin, y los modelos se ajustaron con otro grupo de imágenes de 369 sujetos. En este proceso se construyeron 200 subespacios a partir de muestras del 60 % de los 381 sujetos. Para cada subespacio se ajustaron 20 veces cada tipo de modelo.

Este ejercicio nos permite tener indicadores de rendimiento de los modelos

en condiciones de reproducibilidad de espacios propios de sujetos sanos. Este mismo ejercicio se realizó con los mapas de probabilidad de BIANCA, pero para este caso no se contaba con un conjunto de imágenes solo para la construcción de espacios propios, por lo que se dividió el grupo de los 196 sanos del respectivo conjunto de imágenes.

5.3.1. Eigenfaces por hemisferios y grupo con valores de intensidades en $[0, 1]$: Enfoque A

Los subespacios encontrados en las imágenes se construyeron con 119 y 202 componentes para el hemisferios derecho y el izquierdo respectivamente, y para los patológicos fueron 110 y 148.

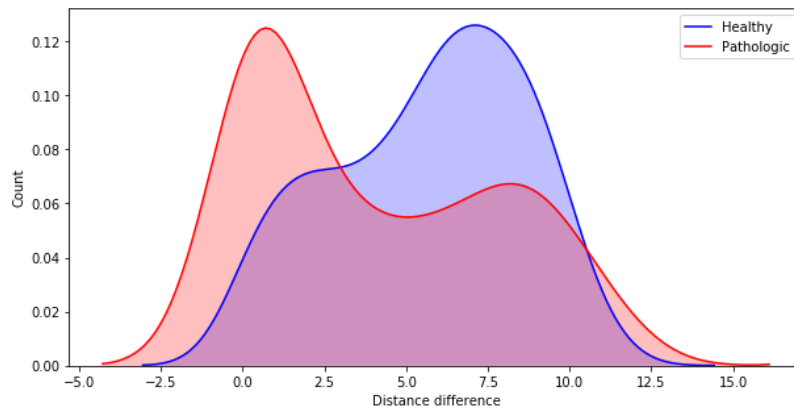


Figura 5.9: Diferencia entre distancia al respectivo espacio propio de cada hemisferios respectivamente.

Las diferencias entre hemisferios de las distancias al respectivo espacio propio muestran distribuciones con formas diferentes, pero con su rango de valores similar; lo que sugiere que la diferencia calculada puede no ser suficiente para distinguir entre ambos grupos (figura 5.9). Para comprobarlo, se ajustaron modelos con las diferencias de las distancias obtenidas.

Después de comparar diferentes modelos de clasificación, las maquinas de soporte vectorial son las que mejor comportamiento tuvieron (5.10). Una vez se hacen predicciones sobre los datos de prueba, se observa un accuracy igual al 61 % (tabla 5.7), lo que en términos generales es un resultado poco satisfactorio para la práctica clínica.

Observe que el desempeño del modelo fue bastante menor en los enfermos, lo que hace que este modelo sea poco viable (5.7).

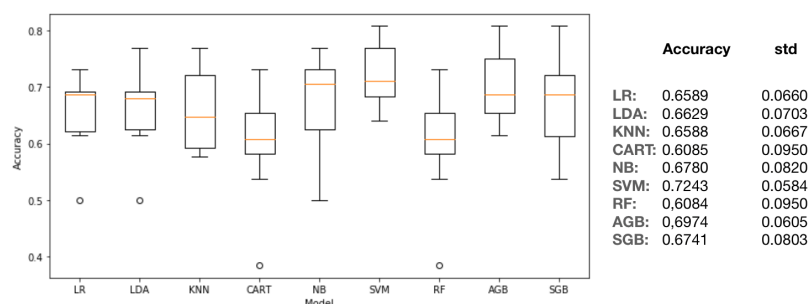


Figura 5.10: Validación cruzada de modelos de clasificación, Eigenfaces enfoque A.

* LR:logístico, LDA: Discriminante lineal, KNN: k-nearest neighbors, CART: Arboles de decisión, NB: Gaussian naïve Bayes, SVM: Máquina de Soporte Vectorial, RF: Random Forest, AGB: AdaBoost, SGB: Gradient Boosted Decision Trees.

Cuadro 5.7: Desempeño del modelo SVM, Eigenfaces enfoque A.

Grupo	Predicción		Indicadores		
	sano	patológico	Precisión	Recall	F1-score
Sanos	51	13	0.63	0.80	0.70
Enfermos	30	17	0.57	0.36	0.44
Total/Promedio	81	30	0.60	0.61	0.59

5.3.2. Eigenfaces con NCC : Enfoque B

Hasta el momento uno de los grandes problemas en las pruebas que se han propuesto, es que la comparación directa entre hemisferios con distancias a subespacios o usando el HCP o el HPLE, no logran captar las características necesarias para hallar diferencias significativas.

Por este motivo se propone trabajar con el NCC completo, es decir, no por hemisferios, sino considerando todo el cerebro. Es importante recordar que el NCC completo es reconstruido a partir de supuestos de simetría entre hemisferios, lo que quiere decir que aunque el método se aplique al cerebro completo, los valores con los que trabaja provienen de medidas de similaridad entre hemisferios.

Antes de continuar se presentan los pasos realizados para los cálculos con el fin de resaltar la importancia que se le dio al abordaje de la reproducibilidad de los resultados a partir de la construcción de espacios propios que provienen de diferentes conjuntos de imágenes:

1. Selección aleatoria del 60 % del conjunto de 381 sujetos sanos.

2. Construcción del espacio propio con Eigenfaces usando las imágenes seleccionadas en el paso 1. Se retiene un 99 % de la variabilidad.
3. Proyección del conjunto de 369 imágenes (212 sanos y 157 patológicos) sobre el espacio propio construido.
4. Validación cruzada (*10-folds*) con los modelos de clasificación propuestos (9 modelos en este caso).
5. Ajuste de cada modelo seleccionado 20 veces (se seleccionan 4 modelos) y cálculo de indicadores para medir rendimiento de modelos.
6. Repetición de los pasos 1 a 5, 200 veces

Observe que cada modelo usado en el paso 5, se corrió 4000 veces con 200 subespacios diferentes. A continuación se muestran los resultados.

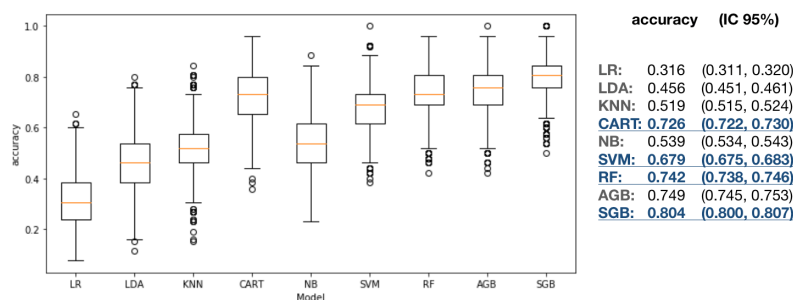


Figura 5.11: Validación cruzada de modelos de clasificación, Eigenfaces enfoque B.

* LR:logístico, LDA: Discriminante lineal, KNN: k-nearest neighbors, CART: Arboles de decisión, NB: Gaussian naïve Bayes, SVM: Máquina de Soporte Vectorial, RF: Random Forest, AGB: AdaBoost, SGB: Gradient Boosted Decision Trees.

Los modelos CART, SVM, RF y SGB son lo que tuvieron un mejor desempeño en la validación cruzada, por esta razón se hace el ajuste con estos cuatro (5.11). Ahora se evaluó el desempeño de los modelos y para el mejor se verifica su capacidad predictiva con el área bajo la curva roc. Observe que hasta el momento el método usado en esta sección es el de mejores resultados.

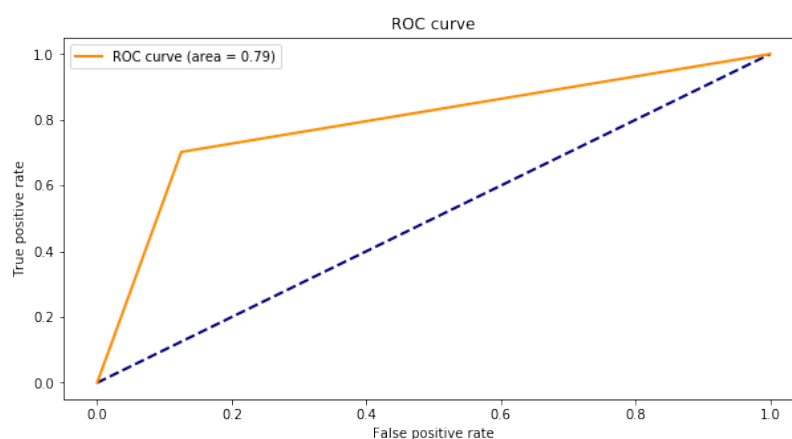
El modelo con mejor rendimiento sobre los datos de prueba fue el SGB con un *accuracy* del 81 %, además es el modelo con el que se consigue una mejor clasificación de sujetos enfermos (tabla 5.8). A continuación se muestran los resultados de ajustar el modelo SGB a partir de las proyecciones realizadas a un espacio propio reteniendo un 90 % de la variabilidad.

Cuadro 5.8: Desempeño de modelos sobre datos de prueba, Eigenfaces enfoque B.

Modelo	Indicador	Sanos (IC 95 %)	Patológicos (IC 95 %)	Accurcy
CART	F1-score	0.768 (0.766,0.769)	0.687 (0.685,0.689)	0.734(0.733,0.736)
	Precision	0.774 (0.772,0.775)	0.687 (0.685,0.689)	
	Recall	0.765 (0.763,0.767)	0.692 (0.690,0.695)	
SVM	F1-score	0.777 (0.776,0.778)	0.534 (0.532,0.536)	0.699 (0.698,0.700)
	Precision	0.678 (0.677,0.679)	0.777 (0.775,0.780)	
	Recall	0.910 (0.908,0.911)	0.411 (0.409,0.413)	
RF	F1-score	0.819 (0.818,0.820)	0.649 (0.646,0.651)	0.761 (0.760,0.763)
	Precision	0.731 (0.730,0.732)	0.856 (0.854,0.858)	
	Recall	0.932 (0.931,0.934)	0.529 (0.526,0.531)	
SGB	F1-score	0.839 (0.837,0.840)	0.782 (0.781,0.784)	0.815 (0.814,0.816)
	Precision	0.845 (0.843,0.846)	0.782 (0.780,0.783)	
	Recall	0.835 (0.834,0.837)	0.787 (0.785,0.789)	

Cuadro 5.9: Desempeño del modelo SGB usando un espacio propio que retiene el 90 % de la variabilidad, Eigenfaces enfoque B.

Grupo	Precision	Recall	F1-score	Accuracy
Helthy	0.80	0.88	0.84	0.8018
Parhologic	0.80	0.70	0.75	

**Figura 5.12:** Capacidad predictiva del modelo SGB usando un espacio propio que retiene el 90 % de la variabilidad, Eigenfaces enfoque B.

El comportamiento del *accuracy* al retener menos variabilidad fue similar, aunque en el *recall* se observó una disminución del 8 % (tabla 5.9). La capacidad predictiva de este modelo fue del 79 % 5.12

5.3.3. Eigenfaces con BIANCA : Enfoque B

En este caso también se generaron 4000 veces cada modelo con 200 subespacios diferentes, con la diferencia que en la validación se usaron 5-folds.

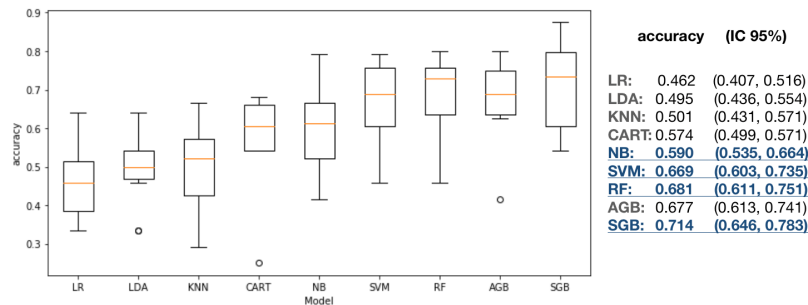


Figura 5.13: Validación cruzada de modelos de clasificación, Eigenfaces enfoque B con BIANCA.

* LR:logístico, LDA: Discriminante lineal, KNN: k-nearest neighbors, CART: Arboles de decisión, NB: Gaussian naïve Bayes, SVM: Máquina de Soporte Vectorial, RF: Random Forest, AGB: AdaBoost, SGB: Gradient Boosted Decision Trees.

Los modelos NB, RF, SVM y SGB son los que mejor comportamiento tuvieron en las pruebas de validación cruzada (figura 5.13). Entre estos cuatro modelos el SGB fue el de mejor comportamiento para detectar sujetos patológicos (tabla 5.10)

Cuadro 5.10: Desempeño de modelos sobre datos de prueba, Eigenfaces enfoque B con BIANCA.

Modelo	Indicador	Sanos (IC 95 %)	Patológicos (IC 95 %)	Accuracy
NB	F1-score	0.663 (0.660,0.665)	0.611 (0.608,0.613)	0.641 (0.639,0.643)
	Precision	0.694 (0.692,0.696)	0.592 (0.589,0.594)	
	Recall	0.641 (0.638,0.645)	0.640 (0.637,0.6436)	
SVM	F1-score	0.747 (0.746,0.749)	0.589 (0.587,0.592)	0.689 (0.687,0.690)
	Precision	0.683 (0.682,0.685)	0.714 (0.711,0.717)	
	Recall	0.829 (0.827,0.832)	0.511 (0.508,0.514)	
RF	F1-score	0.700 (0.672,0.728)	0.585 (0.493,0.676)	0.704 (0.702,0.706)
	Precision	0.683 (0.625,0.740)	0.615 (0.572,0.659)	
	Recall	0.724 (0.676,0.772)	0.565 (0.430,0.700)	
SGB	F1-score	0.732 (0.730,0.734)	0.633 (0.630,0.636)	0.689 (0.687,0.690)
	Precision	0.712 (0.710,0.714)	0.671 (0.668,0.674)	
	Recall	0.758 (0.755,0.761)	0.608 (0.604,0.611)	

El modelo SGB ajustado con datos provenientes de un espacio propio en el que se capturó el 90 % de la variabilidad, logró un *accuracy* y una capacidad predictiva del 66 % (tabla 5.11 y figura 5.14)

Cuadro 5.11: Desempeño del modelo SGB usando un espacio propio que retiene el 90 % de la variabilidad, Eigenfaces enfoque B.

Grupo	Precision	Recall	F1-score	Accuracy
Helthy	0.70	0.58	0.64	0.6596)
Parhologic	0.63	0.74	0.68	

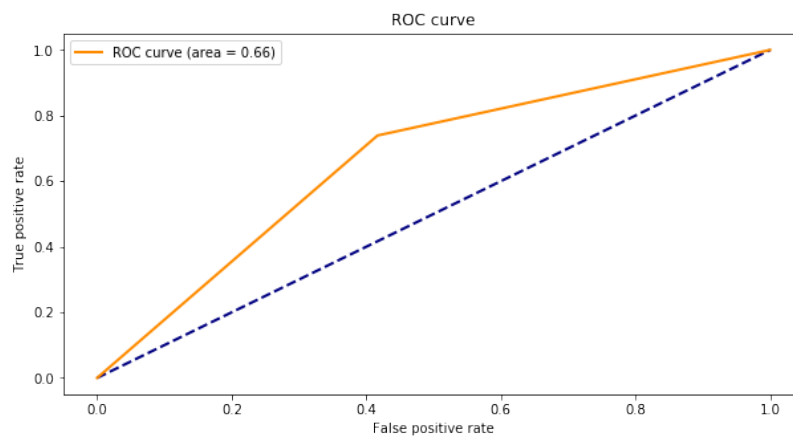


Figura 5.14: Capacidad predictiva del modelo SGB usando un espacio propio que retiene el 90 % de la variabilidad, Eigenfaces enfoque B con Bianca.

Capítulo 6

Discusión y conclusiones

6.1. Discusión

El HCP no logró capturar la suficiente información de las imágenes como para distinguir entre sanos y patológicos, con una detección de patológicos y sanos baja (61 % y 70 % respectivamente) y un *accuracy* del 66 % (tabla 5.2), esto se relaciona con el hecho que los valores del HCP de estos dos grupos solo se diferencia por su varianza y rango (figuras 5.2 y 5.3, tabla 5.2). Es posible que se pueda mejorar estos resultados complementándolo con pruebas de asimetría y curtosis (tercer y cuarto momento), ya que estos dan información adicional sobre la distribución de los datos.

Al aplicar el HCP sobre el mapa de probabilidades calculados con BIANCA, se obtuvieron resultados más pobres que al usar las intensidades en $[0, 1]$, con un 10 % clasificados correctamente para sanos y 77 % para patológicos, con un *accuracy* del 21 % (tabla 5.4); es posible que exista una “perdida” de información al aplicar BIANCA ya que la probabilidad con la que se trabaja esta basado en los $k - \text{vecinos}$ más próximos [19], lo que se podría ver como un tipo de suavizado, aunque no es así técnicamente.

El hecho de que el mapa de probabilidades no funcione muy bien con el HCP, no quiere decir que el método BIANCA arroje pobres resultados, de hecho, se observó un buen comportamiento en la predicción de zonas de hiperintensidades en cerebros patológicas con este método (figura 2.2). En futuros estudios se podría profundizar un poco más tratando de identificar para cuáles tipos o grupos de patologías es más viable usarlo.

Una desventaja del HCP es que es muy sensible a cambios en los criterios o umbrales para definir la pertenencia a un grupo (sección 5.1).

Con el HPLE se observaron mejores resultados para detectar sujetos enfermos comparado con el HCP, con un *accuracy* del 66 % y un *recall* igual al 78 % para un nivel de significancia del 5 % (tabla 5.5). Sin embargo, no es suficiente aun para sugerir que se puede usar en la practica clínica. Una ventaja importante del HPLE sobre el HCP es que se tienen distribuciones por hemisferios de sus valores y se puede complementar con pruebas paramétricas y no paramétricas para determinar si hay diferencias estadísticas (figuras 2.3 y 2.4). Con valores más altos del nivel de significancia se mejora la capacidad para detectar neuroimágenes patológicas del HPLE, pero los demás indicadores de rendimiento disminuyeron drásticamente, es decir que disminuye su capacidad para detectar verdaderos sanos; lo que puede resultar poco ventajoso en la práctica. Un nivel de significancia del 5 % puede ser un valor razonable si se quiere usar el HPLE como medida de detección (figuras 2.3 y 2.4).

Aunque los resultados del HPLE no fueron del todo satisfactorias, es un método interesante que podría ser explorado en futuros trabajos. Con esta medida se logró detectar algunas diferencias usando dos conceptos útiles, el primero es el de permutaciones para aumentar la capacidad de detección de diferencias (disminución del error tipo I) y el segundo tiene que ver con el valor de la entropía como medida de incertidumbre. En otros trabajos se puede probar usando parámetros diferentes para la selección de áreas y proponiendo otras formas de calcular las probabilidades de ocurrencia de intensidades en dichas áreas, también es posible tratar de integrar más información a este indicador o incluso hacer ponderaciones basadas en el grado de intensidad de las áreas seleccionadas.

Finalmente, al aplicar el método de Eigenfaces se obtuvieron los mejores resultados especialmente cuando se trabajó con el mapa de correlaciones NCC en lugar del mapa de probabilidades (figuras 5.12 y 5.14). El modelo con mejor rendimiento y poder predictivo fue el SGB con un 79 % de capacidad predictiva (tabla 5.8 y figura 5.12). En otros trabajos realizados para la identificación de rostros, encontraron que un modelo de maquinas de soporte vectorial proporcionó mejores resultados, pero en este caso los autores usaron imágenes en dos dimensiones [15]. También es posible que las características propias de imágenes de rostros y cerebros influyan en la capacidad predictiva dependiendo del tipo de modelo usado. En este trabajo no se usaron redes neuronales para la clasificación; en futuros estudios en los que se cuente con un mayor número de sujetos se podría probar el rendimiento de este tipo de arquitecturas.

Un aspecto interesante al observar los resultados obtenidos con Eigenfaces

y compararlo con los otros métodos, es que se evidencia que la información adicional capturado por los subespacios (que el HCP y HPLE no consideraron) le permite a los modelos de clasificación tener un mejor desempeño comparado con el HCP y HPLE, además de ser un método computacionalmente viable al hacer una variación en la matriz de covarianzas usada en PCA [48, 57].

Los resultados obtenidos en la sección 5.3.2 además de ser los mejores comparado con los demás métodos, tienen una componente de reproducibilidad ya que se obtuvieron 200 subespacios diferentes y en cada uno de estos espacios propios se ajustó 20 veces cada modelo analizado. En este caso se corrieron 4000 por cada tipo para un total de 16000 modelos ajustados, sin contar con los modelos que fueron necesarios ajustar para la validación cruzada donde se usaron 10-folds para cada tipo (9 modelos de clasificación).

El ejercicio de reproducibilidad realizado en Eigenfaces (secciones 5.3.2 y 5.3.3), permite asegurar que los resultados obtenidos con este método no fueron debidos a la selección de un espacio propio en particular, es decir que los resultados obtenidos se le pueden atribuir al método en sí mismo. Claro está sin desconocer que el uso del NCC como medida de similaridad entre hemisferios permite asegurar que se trabajó con datos libres de sesgos introducidos por parámetros y protocolos en la adquisición de imágenes.

La principal contribución de este trabajo fue la integración del NCC como medidas de similaridad entre hemisferios, con los métodos Eigenfaces y modelos de clasificación, para distinguir entre un cerebro sano y uno enfermo.

Queda por mejorar el HPLE como medida que promete buenos resultados, además de estudiar nuevas aplicaciones con BIANCA para la detección de anomalías en grupos de enfermedades neurodegenerativas o de trastorno mental con posibles afectaciones anatómicas del cerebro.

6.2. Conclusiones

El uso del mapa de correlaciones NCC obtenido de una secuencia FLAIR axial para medir similaridad entre hemisferios resultó ser útil para clasificar sujetos con cerebros sanos o patológicos, después de generar un espacio propio en descomposición de valores singulares con el método Eigenfaces y aplicar modelos de clasificación. Estos métodos aplicados en conjunto, usando un modelo *Stochastic Gradient Boosting* permitieron hacer una clasificación con un poder predictivo de al menos 79 %, dato que podría servir como ayuda para los radiólogos si se quieren tener listas de trabajo priorizadas de manera previa. Esta

priorización puede ser vista en el ámbito clínico como un triaje inicial que ayude a optimizar los flujos de trabajo de los médicos radiólogos en neuroimágenes.

El HPLE es una medida interesante que puede ser estudiada con más detalle integrando otros conceptos y haciendo algunas variaciones tratando de mejorar su capacidad de detección de patrones en imágenes.

Apéndice A

Propuesta de otras líneas de trabajo relacionadas

Se propone trabajar con cuatro medidas que desde el punto de vista clínico y anatómico cobran relevancia y es posible que permitan detectar patologías específicas tales como enfermedades neurodegenerativas, estas medias son: volumen de sustancia gris, volumen de sustancia blanca, volumen del hipocampo e integrarlo con los resultados del volumen de hemisferios presentados en este trabajo.

Las variables propuestas son volúmenes con las que se podrían medir alteraciones anatómicas importantes que están relacionadas con alguna patologías, por ejemplo se sabe que la atrofia de sustancia gris, sustancia blanca y del hipocampo están asociados a trastornos neurológicos [23, 3, 41].

Existe evidencia para apoyar la hipótesis que asegura que la cantidad de sustancia gris y sustancia blanca en el cerebro disminuye con la edad [18, 21], sin embargo una pérdida excesiva de estas se puede asociar a desordenes neurológicos como el Alzheimer, Parkinson, entre otras [41, 10, 55, 33]. Como un criterio o indicador de los volúmenes de estas sustancias es posible establecer valores de referencia según la edad con el fin de detectar posibles anomalías [14].

Resulta interesante incluir un área específica como el hipocampo ya que esta estructura tiene un papel importante en el cerebro por su relación con la memoria y el aprendizaje [27]. Se sabe que alteraciones en esta estructura puede relacionarse también con algunos trastornos neurológicos [25, 36].

Apéndice B

Código

El procesamiento de las imágenes se hizo con FSL mediante rutinas escritas en bash, mientras el procesamiento de datos y la analítica se hizo en Python versión 3.5. El código usado se encuentra disponible en un repositorio de acceso libre en la url https://github.com/styvens/priorizacion_neuroimagenes_RM, donde se incluye también este documento.

En este repositorio se encuentran varios *Notebooks* y *scripts* donde se desarrollan los resultados obtenidos en el proyecto. Se espera que estos resultados sigan siendo alimentados y/o actualizados por personas que trabajen o deseen trabajar en esta línea.

Bibliografía

- [1] Charu C. Aggarwal. *Neural networks and deep learning*. Springer, 2018. ISBN: 3-319-94463-0.
- [2] José Francisco Camacho Aguilera. «Charcot y su Legado a la Medicina». ES. En: *Gaceta Médica de México* 148.3 (2012), págs. 321-326. ISSN: 0016-3813. URL: <https://www.medigraphic.com/cgi-bin/new/resumen.cgi?IDARTICULO=35565&id2=> (visitado 20-11-2019).
- [3] Houshang Amiri y col. «Urgent challenges in quantification and interpretation of brain grey matter atrophy in individual MS patients using MRI». En: *NeuroImage : Clinical* 19 (abr. de 2018), págs. 466-475. ISSN: 2213-1582. DOI: [10.1016/j.nicl.2018.04.023](https://doi.org/10.1016/j.nicl.2018.04.023). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6030805/> (visitado 12-11-2019).
- [4] Bouzalmat Anissa y col. «Face recognition: Comparative study between linear and non linear dimensionality reduction methods». En: *2015 International Conference on Electrical and Information Technologies (ICEIT)*. Mar. de 2015, págs. 224-228. DOI: [10.1109/EITech.2015.7162932](https://doi.org/10.1109/EITech.2015.7162932).
- [5] Syed Muhammad Anwar y col. «Medical Image Analysis using Convolutional Neural Networks: A Review». en. En: *Journal of Medical Systems* 42.11 (oct. de 2018), pág. 226. ISSN: 1573-689X. DOI: [10.1007/s10916-018-1088-1](https://doi.org/10.1007/s10916-018-1088-1). URL: <https://doi.org/10.1007/s10916-018-1088-1> (visitado 08-07-2020).
- [6] Axel Bacher y col. «Generating Random Permutations by Coin Tossing: Classical Algorithms, New Analysis, and Modern Implementation». En: *ACM Transactions on Algorithms* 13.2 (feb. de 2017), 24:1-24:43. ISSN: 1549-6325. DOI: [10.1145/3009909](https://doi.org/10.1145/3009909). URL: <https://doi.org/10.1145/3009909> (visitado 19-06-2020).
- [7] Tom L. Beauchamp. «The belmont report». En: *The Oxford textbook of clinical research ethics* (2008). Publisher: Oxford University Press Oxford, págs. 149-155.
- [8] Matthew Brett, Ingrid S. Johnsrude y Adrian M. Owen. «The problem of functional localization in the human brain». eng. En: *Nature Reviews. Neuroscience* 3.3 (mar. de 2002), págs. 243-249. ISSN: 1471-003X. DOI: [10.1038/nrn756](https://doi.org/10.1038/nrn756).

- [9] Kerstin Bunte y col. «Stochastic neighbor embedding (SNE) for dimension reduction and visualization using arbitrary divergences». en: *Neurocomputing*. Advances in artificial neural networks, machine learning, and computational intelligence (ESANN 2011) 90 (ago. de 2012), págs. 23-45. ISSN: 0925-2312. DOI: [10.1016/j.neucom.2012.02.034](https://doi.org/10.1016/j.neucom.2012.02.034). URL: <http://www.sciencedirect.com/science/article/pii/S0925231212001920> (visitado 10-07-2020).
- [10] Fu-Xiang Chen y col. «Gray matter atrophy associated with mild cognitive impairment in Parkinson's disease». eng. En: *Neuroscience Letters* 617 (mar. de 2016), págs. 160-165. ISSN: 1872-7972. DOI: [10.1016/j.neulet.2015.12.055](https://doi.org/10.1016/j.neulet.2015.12.055).
- [11] W. Conover. «Two Independent Samples». En: *Practical Nonparametric Statistics*. New York: John Wiley&Sons Inc., 1999, págs. 271-288.
- [12] Gloria Díaz y col. «Automatic classification of structural mri for diagnosis of neurodegenerative diseases». En: *Acta Biológica Colombiana* 15.3 (dic. de 2010). Publisher: Universidad Nacional de Colombia, págs. 165-180. ISSN: 0120-548X. URL: http://www.scielo.org.co/scielo.php?script=sci_abstract&pid=S0120-548X2010000300012&lng=en&nrm=iso&tlng=en (visitado 08-07-2020).
- [13] Michael Falk. «On Mad and Comedians». en: En: *Annals of the Institute of Statistical Mathematics* 49.4 (dic. de 1997), págs. 615-644. ISSN: 1572-9052. DOI: [10.1023/A:1003258024248](https://doi.org/10.1023/A:1003258024248). URL: <https://doi.org/10.1023/A:1003258024248> (visitado 22-06-2020).
- [14] Nikdokht Farid y col. «Temporal Lobe Epilepsy: Quantitative MR Volumetry in Detection of Hippocampal Atrophy». En: *Radiology* 264.2 (ago. de 2012), págs. 542-550. ISSN: 0033-8419. DOI: [10.1148/radiol.12112638](https://doi.org/10.1148/radiol.12112638). URL: <https://pubs.rsna.org/doi/full/10.1148/radiol.12112638> (visitado 19-11-2019).
- [15] Md. Omar Faruque y Md. Al Mehedi Hasan. «Face recognition using PCA and SVM». En: *and Identification in Communication 2009 3rd International Conference on Anti-counterfeiting, Security*. ISSN: 2163-5056. Ago. de 2009, págs. 97-101. DOI: [10.1109/ICASID.2009.5276938](https://doi.org/10.1109/ICASID.2009.5276938).
- [16] Pauline A. Filipek y col. «The Young Adult Human Brain: An MRI-based Morphometric Analysis». en: En: *Cerebral Cortex* 4.4 (jul. de 1994), págs. 344-360. ISSN: 1047-3211. DOI: [10.1093/cercor/4.4.344](https://doi.org/10.1093/cercor/4.4.344). URL: <https://academic.oup.com/cercor/article/4/4/344/323479> (visitado 29-10-2019).
- [17] Jerome Friedman, Trevor Hastie y Robert Tibshirani. «Unsupervised Learning». En: *The elements of statistical learning*. Vol. 1. Issue: 10. Springer series in statistics New York, 2001, págs. 534-550.

- [18] Yulin Ge y col. «Age-Related Total Gray Matter and White Matter Changes in Normal Adult Brain. Part I: Volumetric MR Imaging Analysis». en. En: *American Journal of Neuroradiology* 23.8 (sep. de 2002), págs. 1327-1333. ISSN: 0195-6108, 1936-959X. URL: <http://www.ajnr.org/content/23/8/1327> (visitado 19-11-2019).
- [19] Ludovica Griffanti y col. «BIANCA (Brain Intensity AbNormality Classification Algorithm): A new tool for automated segmentation of white matter hyperintensities». en. En: *NeuroImage* 141 (nov. de 2016), págs. 191-205. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2016.07.018](https://doi.org/10.1016/j.neuroimage.2016.07.018). URL: <http://www.sciencedirect.com/science/article/pii/S1053811916303251> (visitado 23-02-2020).
- [20] Manjari Gupta y col. «Design Pattern Detection by normalized cross correlation». En: *2010 International Conference on Methods and Models in Computer Science (ICM2CS-2010)*. Dic. de 2010, págs. 81-84. DOI: [10.1109/ICM2CS.2010.5706723](https://doi.org/10.1109/ICM2CS.2010.5706723).
- [21] C.R.G. Guttmann y col. «White matter changes with normal aging». En: *Neurology* 50.4 (abr. de 1998), pág. 972. DOI: [10.1212/WNL.50.4.972](https://doi.org/10.1212/WNL.50.4.972). URL: <http://n.neurology.org/content/50/4/972.abstract>.
- [22] Steve Hanneke. «The Optimal Sample Complexity of PAC Learning». En: *arXiv:1507.00473 [cs, stat]* (feb. de 2016). arXiv: 1507.00473. URL: <http://arxiv.org/abs/1507.00473> (visitado 16-07-2020).
- [23] Mamoru Hashimoto y col. «Does donepezil treatment slow the progression of hippocampal atrophy in patients with Alzheimer's disease?» eng. En: *The American Journal of Psychiatry* 162.4 (abr. de 2005), págs. 676-682. ISSN: 0002-953X. DOI: [10.1176/appi.ajp.162.4.676](https://doi.org/10.1176/appi.ajp.162.4.676).
- [24] Trevor Hastie, Robert Tibshirani y Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009. ISBN: 0-387-84858-4.
- [25] R. Horn y col. «Atrophy of Hippocampus in Patients with Alzheimer's Disease and Other Diseases with Memory Impairment». Inglés. En: *Dementia and Geriatric Cognitive Disorders* 7.4 (1996), págs. 182-186. ISSN: 1420-8008, 1421-9824. DOI: [10.1159/000106876](https://doi.org/10.1159/000106876). URL: <https://www.karger.com/Article/FullText/106876> (visitado 19-11-2019).
- [26] Laura Igual y Santi Seguí. «Supervised Learning». En: *Introduction to Data Science*. Springer, 2017, págs. 67-96.
- [27] Daniel Johnston y David G. Amaral. «Hippocampus». en_US. En: *The Synaptic Organization of the Brain*. Fifth. Oxford University Press, ene. de 2004, págs. 455-498. ISBN: 978-0-19-986444-7. (Visitado 19-11-2019).

- [28] Ian T. Jolliffe y Jorge Cadima. «Principal component analysis: a review and recent developments». En: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374.2065 (abr. de 2016). Publisher: Royal Society, pág. 20150202. DOI: [10.1098/rsta.2015.0202](https://royalsocietypublishing.org/doi/10.1098/rsta.2015.0202). URL: <https://royalsocietypublishing.org/doi/10.1098/rsta.2015.0202> (visitado 10-07-2020).
- [29] Margarida Julià-Sapé y col. «Comparison between neuroimaging classifications and histopathological diagnoses using an international multicenter brain tumor magnetic resonance imaging database». en_US. En: *Journal of Neurosurgery* 105.1 (jul. de 2006). Publisher: American Association of Neurological Surgeons Section: Journal of Neurosurgery, págs. 6-14. DOI: [10.3171/jns.2006.105.1.6](https://thejns.org/view/journals/j-neurosurg/105/1/article-p6.xml). URL: <https://thejns.org/view/journals/j-neurosurg/105/1/article-p6.xml> (visitado 08-07-2020).
- [30] P. Kalavathi, M. Senthamilselvi y V. B. Surya Prasath. «Review of Computational Methods on Brain Symmetric and Asymmetric Analysis from Neuroimaging Techniques». en. En: *Technologies* 5.2 (jun. de 2017), pág. 16. DOI: [10.3390/technologies5020016](https://www.mdpi.com/2227-7080/5/2/16). URL: <https://www.mdpi.com/2227-7080/5/2/16> (visitado 30-10-2019).
- [31] Stefan Klöppel y col. «Diagnostic neuroimaging across diseases». en. En: *NeuroImage*. Neuroimaging: Then, now and the future 61.2 (jun. de 2012), págs. 457-463. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2011.11.002](http://www.sciencedirect.com/science/article/pii/S105381191101281X). URL: <http://www.sciencedirect.com/science/article/pii/S105381191101281X> (visitado 08-07-2020).
- [32] Max Kuhn y Kjell Johnson. «Classification Models». En: *Applied predictive modeling*. Vol. 26. Springer, 2013, págs. 247-457.
- [33] Jonas Kunst y col. «Patterns of Grey Matter Atrophy at Different Stages of Parkinson's and Alzheimer's Diseases and Relation to Cognition». eng. En: *Brain Topography* 32.1 (2019), págs. 142-160. ISSN: 1573-6792. DOI: [10.1007/s10548-018-0675-2](https://doi.org/10.1007/s10548-018-0675-2).
- [34] Vadim Kuperman. «Introduction». En: *Magnetic Resonance Imaging*. United States of America: Academic Press, 2000, págs. 1-7.
- [35] Pradyumna Lanka y col. «Supervised machine learning for diagnostic classification from large-scale neuroimaging datasets». en. En: *Brain Imaging and Behavior* (nov. de 2019). ISSN: 1931-7565. DOI: [10.1007/s11682-019-00191-8](https://doi.org/10.1007/s11682-019-00191-8). URL: <https://doi.org/10.1007/s11682-019-00191-8> (visitado 08-07-2020).
- [36] Kangyoon Lee y col. «Right hippocampus atrophy is independently associated with Alzheimer's disease with psychosis». eng. En: *Psychogeriatrics*:

- The Official Journal of the Japanese Psychogeriatric Society* 19.2 (mar. de 2019), págs. 105-110. ISSN: 1479-8301. DOI: [10.1111/psyg.12369](https://doi.org/10.1111/psyg.12369).
- [37] Weiming Lin y col. «Convolutional Neural Networks-Based MRI Image Analysis for the Alzheimer's Disease Prediction From Mild Cognitive Impairment». English. En: *Frontiers in Neuroscience* 12 (2018). Publisher: Frontiers. ISSN: 1662-453X. DOI: [10.3389/fnins.2018.00777](https://doi.org/10.3389/fnins.2018.00777). URL: <https://www.frontiersin.org/articles/10.3389/fnins.2018.00777/full> (visitado 08-07-2020).
- [38] Sheena Xin Liu. «Symmetry and asymmetry analysis and its implications to computer-aided diagnosis: A review of the literature». en. En: *Journal of Biomedical Informatics* 42.6 (dic. de 2009), págs. 1056-1064. ISSN: 1532-0464. DOI: [10.1016/j.jbi.2009.07.003](https://doi.org/10.1016/j.jbi.2009.07.003). URL: <http://www.sciencedirect.com/science/article/pii/S1532046409000975> (visitado 30-10-2019).
- [39] J. M. Martí-Climent y col. «[Neuroimaging: technical aspects and practice]». spa. En: *Revista Espanola De Medicina Nuclear* 29.4 (ago. de 2010), págs. 189-210. ISSN: 0212-6982. DOI: [10.1016/j.rem.2010.05.001](https://doi.org/10.1016/j.rem.2010.05.001).
- [40] Código de Núremberg. «Tribunal Internacional de Núremberg». En: *Recuperado de* http://www.sld.cu/galerias/pdf/sitios/prevemi/codigo_nuremberg.pdf (1947).
- [41] Miklos Palotai y col. «History of fatigue in multiple sclerosis is associated with grey matter atrophy». eng. En: *Scientific Reports* 9.1 (oct. de 2019), pág. 14781. ISSN: 2045-2322. DOI: [10.1038/s41598-019-51110-2](https://doi.org/10.1038/s41598-019-51110-2).
- [42] Saima Rathore y col. «A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages». En: *NeuroImage* 155 (jul. de 2017), págs. 530-548. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2017.03.057](https://doi.org/10.1016/j.neuroimage.2017.03.057). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5511557/> (visitado 08-07-2020).
- [43] C. Rottschy y col. «Modelling neural correlates of working memory: A coordinate-based meta-analysis». En: *NeuroImage* 60.1 (mar. de 2012), págs. 830-846. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2011.11.050](https://doi.org/10.1016/j.neuroimage.2011.11.050). URL: <http://www.sciencedirect.com/science/article/pii/S1053811911013310> (visitado 15-10-2019).
- [44] Clément W. Royer, Michael O'Neill y Stephen J. Wright. «A Newton-CG Algorithm with Complexity Guarantees for Smooth Unconstrained Optimization». En: *arXiv:1803.02924 [math]* (nov. de 2018). arXiv: 1803.02924. URL: <http://arxiv.org/abs/1803.02924> (visitado 18-07-2020).
- [45] Andrea del Pilar Rueda-O. y col. «Una revisión de técnicas básicas de neuroimagen para el diagnóstico de enfermedades neurodegenerativas». En: *Biosalud* 17.2 (jul. de 2018). Publisher: Universidad de Caldas, págs. 59-90.

- ISSN: 1657-9550. DOI: [10.17151/biosa.2018.17.2.5](https://doi.org/10.17151/biosa.2018.17.2.5). URL: http://www.scielo.org.co/scielo.php?script=sci_abstract&pid=S1657-95502018000200059&lng=es&nrm=iso&tlng=es (visitado 08-07-2020).
- [46] P. Sartori y col. «Artefactos y artificios frecuentes en tomografía computada y resonancia magnética». es. En: *Revista Argentina de Radiología* 79.4 (oct. de 2015), págs. 192-204. ISSN: 0048-7619. DOI: [10.1016/j.rard.2015.04.005](https://doi.org/10.1016/j.rard.2015.04.005). URL: <http://www.sciencedirect.com/science/article/pii/S0048761915000794> (visitado 16-07-2020).
- [47] C. E. Shannon. «A mathematical theory of communication». En: *ACM SIGMOBILE Mobile Computing and Communications Review* 5.1 (ene. de 2001), págs. 3-55. ISSN: 1559-1662. DOI: [10.1145/584091.584093](https://doi.org/10.1145/584091.584093). URL: <https://doi.org/10.1145/584091.584093> (visitado 21-06-2020).
- [48] L. Sirovich y M. Kirby. «Low-dimensional procedure for the characterization of human faces». EN. En: *JOSA A* 4.3 (mar. de 1987), págs. 519-524. ISSN: 1520-8532. DOI: [10.1364/JOSAA.4.000519](https://doi.org/10.1364/JOSAA.4.000519). URL: <https://www.osapublishing.org/josaa/abstract.cfm?uri=josaa-4-3-519> (visitado 11-11-2019).
- [49] Steven S. Skiena. «Machine Learning». En: *The data science design manual*. Springer, 2017, págs. 351-371. ISBN: 3-319-55444-1.
- [50] Stephen M. Smith. «Fast robust automated brain extraction». eng. En: *Human Brain Mapping* 17.3 (nov. de 2002), págs. 143-155. ISSN: 1065-9471. DOI: [10.1002/hbm.10062](https://doi.org/10.1002/hbm.10062).
- [51] Benjamín B. Spota. «Histórico y evolución de la neurología». es. En: *Archivos de Neuro-Psiquiatria* 4.3 (sep. de 1946), págs. 295-303. ISSN: 0004-282X. DOI: [10.1590/S0004-282X1946000300005](https://doi.org/10.1590/S0004-282X1946000300005). URL: http://www.scielo.br/scielo.php?script=sci_abstract&pid=S0004-282X1946000300005&lng=en&nrm=iso&tlng=es (visitado 20-11-2019).
- [52] Taylor J. Stinnett y Matthew K. Zabel. «Neuroanatomy, Broca Area». eng. En: *StatPearls*. Treasure Island (FL): StatPearls Publishing, 2019. URL: <http://www.ncbi.nlm.nih.gov/books/NBK526096/> (visitado 29-10-2019).
- [53] Li Sun y col. «Brain Tumor Segmentation and Survival Prediction Using Multimodal MRI Scans With Deep Learning». English. En: *Frontiers in Neuroscience* 13 (2019). ISSN: 1662-453X. DOI: [10.3389/fnins.2019.00810](https://doi.org/10.3389/fnins.2019.00810). URL: <https://www.frontiersin.org/articles/10.3389/fnins.2019.00810/full> (visitado 29-10-2019).
- [54] M. Symms y col. «A review of structural magnetic resonance neuroimaging». en. En: *Journal of Neurology, Neurosurgery & Psychiatry* 75.9 (sep. de 2004), págs. 1235-1244. ISSN: 0022-3050, 1468-330X. DOI: [10.1136/jnnp](https://doi.org/10.1136/jnnp).

- 2003.032714. URL: <https://jnnp.bmj.com/content/75/9/1235> (visitado 28-10-2019).
- [55] Windsor Kwan-Chun Ting y col. «Grey matter atrophy in mild cognitive impairment / early Alzheimer disease associated with delusions: a voxel-based morphometry study». eng. En: *Current Alzheimer Research* 12.2 (2015), págs. 165-172. ISSN: 1875-5828. DOI: [10.2174/1567205012666150204130456](https://doi.org/10.2174/1567205012666150204130456).
- [56] Arthur W. Toga y Paul M. Thompson. «Mapping brain asymmetry». eng. En: *Nature Reviews. Neuroscience* 4.1 (ene. de 2003), págs. 37-48. ISSN: 1471-003X. DOI: [10.1038/nrn1009](https://doi.org/10.1038/nrn1009).
- [57] M.A. Turk y A.P. Pentland. «Face recognition using eigenfaces». En: *1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Proceedings*. ISSN: 1063-6919. Jun. de 1991, págs. 586-591. DOI: [10.1109/CVPR.1991.139758](https://doi.org/10.1109/CVPR.1991.139758).
- [58] Matthew Turk y Alex Pentland. «Eigenfaces for recognition». En: *Journal of Cognitive Neuroscience* 3.1 (ene. de 1991), págs. 71-86. ISSN: 0898-929X. DOI: [10.1162/jocn.1991.3.1.71](https://doi.org/10.1162/jocn.1991.3.1.71). URL: <https://doi.org/10.1162/jocn.1991.3.1.71> (visitado 22-06-2020).
- [59] Javier E. Villanueva-Meyer, Marc C. Mabray y Soonmee Cha. «Current Clinical Brain Tumor Imaging». En: *Neurosurgery* 81.3 (sep. de 2017), págs. 397-415. ISSN: 0148-396X. DOI: [10.1093/neuros/nyx103](https://doi.org/10.1093/neuros/nyx103). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5581219/> (visitado 08-07-2020).
- [60] Li Xiang, Timothy Crow y Neil Roberts. «Cerebral torque is human specific and unrelated to brain size». en. En: *Brain Structure and Function* 224.3 (abr. de 2019), págs. 1141-1150. ISSN: 1863-2661. DOI: [10.1007/s00429-018-01818-0](https://doi.org/10.1007/s00429-018-01818-0). URL: <https://doi.org/10.1007/s00429-018-01818-0> (visitado 29-10-2019).